Mindfulness and Compassion: Measurement and Mechanisms of Interventions

Jenny Gu

Thesis submitted for the degree of Doctor of Philosophy

School of Psychology

University of Sussex

May 2018

Declaration

This thesis conforms to an 'article format', whereby Chapters 2 to 7 are distinct papers written in a style appropriate for publication in peer-reviewed journals. Chapter 1 gives an overview of mindfulness and compassion theory and research, and Chapter 8 presents a general discussion of the research undertaken. Tables and figures are presented in numerical order across the thesis. Full references of citations in each chapter are presented together at the end of the thesis.

Chapters 2 to 6 have been published in peer-reviewed journals and Chapter 7 is under peer review; references for these papers and author contributions are given on the title pages for these chapters.

I hereby declare that this thesis has not been, and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature: ……

Jenny Gu

Date: ……

Acknowledgements

My deepest gratitude to the following people for their role in shaping my PhD thesis and experience:

First of all, to my supervisors, Dr Clara Strauss and Dr Kate Cavanagh, for being wonderful role models and mentors, for believing in me and providing me with many opportunities to grow as a researcher. I couldn't have been better guided through my PhD.

I'm lucky to have been able to pick the brains of not one, but two stats whizzes. Prof Andy Field – thank you for your helpful advice as my internal assessor and Dr Rod Bond – thank you for patiently answering my endless SEM questions.

I'm also grateful to have worked with, and learnt from, the brilliant Prof Ruth Baer and Prof Willem Kuyken. Thank you both for encouraging and inspiring me.

A special thanks to my lovely PhD friends and colleagues, in particular Moitree Banerjee, Heather Taylor, James Moon, Sarah Becker, and Anne Templeton. A quote comes to mind: "Friendship is unnecessary, like philosophy, like art. . . . It has no survival value; rather it is one of those things that gives value to survival." - C. S. Lewis. Thank you all for giving value to this experience.

To my family and first teachers – thank you for your quiet and unwavering love and support. Finally, to Joe, my partner in life and best friend, I'm forever grateful for all that you do. For your love and companionship, for providing comic relief, for supporting me in all my endeavours – the list goes on.

UNIVERSITY OF SUSSEX

Jenny Gu

Thesis submitted for the degree of Doctor of Philosophy

Mindfulness and Compassion: Measurement and Mechanisms of Interventions

Summary

In recent years, there has been an exponential increase in research exploring contemplative constructs, namely mindfulness and compassion, and their potential to enhance psychological functioning. A large body of evidence supports the effectiveness of mindfulness-based interventions (MBIs) for improving mental health and wellbeing, and emerging evidence indicates benefits associated with cultivating compassion. However, significant gaps remain, which impede progress in mindfulness and compassion research. Understanding and empirical testing of the mechanisms underlying the effects of MBIs are limited. Research on MBIs also requires valid and reliable mindfulness measures and existing self-report scales need additional psychometric testing. Despite increasing research attention on self- and other-compassion, there is a lack of definitional clarity and psychometrically robust measures of these constructs. This thesis aims to address these omissions.

Following an overview of mindfulness and compassion theory and research, Chapter 2 presents a systematic review and meta-analysis of studies which formally tested mechanisms of MBIs. Chapter 3 examines the specific effects and mechanisms of learning mindfulness, by comparing an online self-help MBI with a matched control condition. Much of effectiveness and mechanism research involves comparing mindfulness scores before and after MBIs, yet the factor structure of commonly used self-report measures before and after MBIs has not been tested; this is addressed in Chapter 4. Chapters 5, 6, and 7 focus on increasing clarity in defining and measuring compassion. Chapter 5 proposes a five-element definition of compassion and includes a systematic review and evaluation of existing compassion measures. Chapter 6 empirically tests the factor structure of the five-element definition using self-report items. Following theoretical and empirical support for the five-element conceptualisation, Chapter 7 uses this definition to develop and validate new self-report measures of self- and other-compassion. Chapter 8 presents a general discussion of the

research undertaken, including strengths and limitations, future directions, and implications.

Table of Contents

List of Tables

List of Figures

Chapter 1:
General Introduction

Chapter Overview

Over the last few decades, scientific and societal interest in contemplative constructs and practices has grown rapidly. Whilst research on mindfulness dominates the field, compassion research is receiving increased attention. In contemplative traditions, the cultivation of mindfulness and compassion are intimately linked (e.g., Rapgay & Bystrisky, 2009; Thurman, 1997) and empirical studies have found these constructs to be strongly related (e.g., Gu, Baer, Cavanagh, Kuyken, & Strauss, under review).

Research has focused on the potential of psychological interventions grounded in mindfulness and compassion principles for improving mental health and wellbeing. The prevalence and impact of poor mental health is well documented; depression is the leading contributor to global disability of any health condition (physical or mental) and anxiety disorders are the sixth leading contributor (World Health Organisation, 2017). In addition to the personal impact they have, depression and anxiety disorders are estimated to cost the UK economy £12.15 billion and £14.19 billion per year by 2026, respectively (McCrone, Dhanasiri, Patel, Knapp, & Lawton-Smith, 2008). There is also increasing recognition of the prevalence and negative impact of stress and poor wellbeing in the general population (e.g., Chartered Institute of Personnel and Development, 2016; Hassard, Teoh, Visockaite, Dewe, & Cox, 2018; Health and Safety Executive, 2017; Mental Health Foundation, 2018).

Evidence shows that mindfulness-based interventions (MBIs) and compassion-based interventions (CBIs) could be part of the solution to treating poor mental health and wellbeing. Meta-analyses of randomised controlled trials (RCTs) have shown that MBIs are effective for reducing the relative risk of depressive relapse (Kuyken et al., 2016), for leading to lower levels of depression symptom severity for people experiencing a current episode of depression (Strauss, Cavanagh, Oliver, & Pettman, 2014), and for leading to improved stress and wellbeing outcomes (Khoury, Sharma, Rush, & Fournier, 2015). Similarly, a recent meta-analysis of RCTs for CBIs indicates that they lead to lower levels of depression, anxiety, wellbeing, and psychological distress (Kirby, Tellegen, & Steindl, 2017).

To ensure the rigor of the evidence base for MBIs and CBIs and thereby increase their public health impact, several research areas require attention: we need to

ensure that the constructs of mindfulness and compassion are clearly defined and have reliable and valid measures, and MBIs and CBIs need to be robustly tested to determine whether they work through their proposed mechanisms. These key areas are addressed in this programme of research with the hope that this will contribute to improved mental health care.

The following sections in this General Introduction chapter present an overview of first, mindfulness theory and research, and next, compassion theory and research, ending with the aims of the current thesis and overview of chapters. Given that research on mindfulness and compassion are at different stages of development, the below sections in this chapter on the two contemplative constructs do not follow the same structure (e.g., mechanism theory and research are discussed for mindfulness interventions but not compassion interventions).

What is Mindfulness?

> "Let any one try, I will not say to arrest, but to notice or attend to, the present moment of time. One of the most baffling experiences occurs. Where is it, this present? It has melted in our grasp, fled ere we could touch it, gone in the instant of becoming."
> – William James

Mindfulness originates from Buddhist traditions and is translated from the Pali term *sati*, meaning awareness, attention, and remembrance (Siegel, Germer, & Olendzki, 2009). Awareness refers to open, continuous observation of present moment experience, attention indicates a deep and calm focus on a restricted field of experience, and remembrance is the continued intention to maintain awareness and attention. The term *sati* is best translated as a process ('to be mindful'), which reflects the importance given in Buddhism to maintaining a regular, experiential meditation practice in order to first refine attention and awareness and ultimately gain insight and wisdom (Grossman & Van Dam, 2011). In this way, mindfulness is regarded as essential for managing the mind (Bhikkhu, 2007) and provides the foundation for the development of other positive mental capacities (e.g., compassion) (Rapgay & Bystrisky, 2009).

There are many definitions of mindfulness in psychological literature, which are for the most part consistent with Buddhist conceptualisations (Bishop, 2002) and typically describe two elements (Baer, 2015): *what* it involves (i.e., being aware of and

attentive to present moment experience) and *how* one does it (i.e., the qualities of attention). In relation to the *how* aspect, Baer (2015) notes that definitions of mindfulness generally agree that it is accepting, non-judgemental, open, curious, and compassionate. In this sense, mindfulness entails being aware of and open to the full spectrum of experience in each moment, without judgment, avoidance, or fixation. Kabat-Zinn's (2003) well known definition of mindfulness captures both of these elements; mindfulness is "the awareness that emerges through paying attention on purpose, in the present moment, and non-judgmentally to the unfolding experience moment by moment" (p. 145).

Despite its Buddhist origins, mindfulness is regarded as an inherent and universal human capacity (Kabat-Zinn, 2003). In particular, it is considered to be both an immediate state and a dispositional trait. We are all mindful to a certain degree, moment by moment, and this state fluctuates across time and situations and can be cultivated through mindfulness meditation practice. Dispositional mindfulness refers to our levels of mindfulness which are broadly consistent across time and contexts. This state-trait distinction is supported by recent findings showing that individual trajectories of mindfulness states cultivated through meditation predict increases in trait mindfulness (Kiken, Garland, Bluth, Palsson, & Gaylord, 2015).

Mindfulness-Based Interventions

In the last few decades, there has been an exponential growth in interest in mindfulness in science and wider society, largely attributed to the development of secular psychological interventions based on mindfulness practice (mindfulness-based interventions; MBIs) which aim to increase mindfulness and as a consequence, improve mental health and wellbeing. Although several interventions incorporate teaching of mindfulness principles and practice, mindfulness-based stress reduction (MBSR; Kabat-Zinn, 1982) and mindfulness-based cognitive therapy (MBCT; Segal, Williams, & Teasdale, 2002; 2013) are the two most widely applied and evaluated MBIs and can be distinguished by their predominant focus on mindfulness practices, their course structure, and the training and accreditation required of intervention facilitators.

Both MBSR and MBCT are manualised group-based courses, consisting of eight consecutive weekly 2 to 2.5-hour sessions, which teach participants mindfulness

skills primarily through engagement in formal and informal mindfulness meditation practices during sessions and at home. Overall capacity of mindful awareness is developed over the course of these programmes, with the development of early skills (e.g., focusing and stabilisation of attention) providing an important foundation for the acquisition of later skills (e.g., learning to relate to difficult experiences in a non-judgemental and non-reactive way).

MBSR was originally designed for the management of chronic pain and stress and has been extended to improve other outcomes. Several meta-analyses have found that compared to control conditions, MBSR is effective for improving a range of outcomes in both people with chronic conditions and healthy individuals, including depression, anxiety, stress, distress, and quality of life, with between-group effect sizes ranging from small to large (e.g., Bohlmeijer, Prenger, Taal, & Cuijpers, 2010; Grossman, Niemann, Schmidt, & Walach, 2004; Khoury et al., 2015).

MBCT was originally developed as a relapse prevention intervention for people with recurrent depression in remission and incorporates MBSR teachings with cognitive behavioural therapy (CBT) for depression strategies, with CBT being an already well-established treatment for current depression. MBCT was developed based on the differential activation hypothesis (DAH; Teasdale, 1988), a model of cognitive vulnerability to depressive relapse which posits that repeated associations between low mood and repetitive negative thinking patterns experienced during early episodes of depression increase the future likelihood that mild dysphoria will reactivate negative thinking patterns, and thus vulnerability to subsequent episodes. MBCT is theorised to decrease depressive relapse by increasing mindful awareness of and disengagement from internal events that have the potential to escalate into an episode of depression (e.g., rumination).

As with MBSR, MBCT has been adapted for other populations (e.g., ‘MBCT for Life’ or MBCT-L adapted for non-clinical populations; Bernard, Cullen, & Kuyken, 2017) and to target other outcomes such as current depressive symptoms, anxiety, stress, and wellbeing. This is theoretically justified, as repetitive negative thinking processes (rumination and worry) are implicated in the onset and maintenance of a broad range of mental health difficulties (e.g., anxiety disorders, major depressive disorder, insomnia, eating disorders, substance abuse) and developing the skills to non-judgementally notice and disengage from rumination and worry are likely to be of therapeutic value. Meta-analyses support this suggestion – they have found that

compared to control conditions, MBCT is effective for reducing the relative risk of depressive relapse (Kuyken et al., 2016; Piet & Hougaard, 2011) and symptom severity in those diagnosed with depressive and anxiety disorders (Strauss et al., 2014). An RCT has also found MBCT-L to be effective compared to a waitlist control group in improving stress, depression, anxiety, and wellbeing outcomes in a healthcare staff sample (Strauss, Gu, Pitman, Chapman, Kuyken, & Whittington, in preparation).

Research has additionally found an association between the amount of formal mindfulness practice at home and risk of depressive relapse; for participants undertaking MBCT, those who engaged in formal home practice on at least 3 days a week were almost half as likely to relapse compared to those who engaged in fewer days of formal home practice (Crane et al., 2014). This supports the premise of MBIs, which regard practice over the course of the programme as essential for the development of mindful awareness and metacognitive insight (*experiencing* thoughts *as* events in the mind) (Teasdale, 1999) and the acquisition of therapeutic benefits. This also supports the emphasis given in Buddhism on maintaining a regular meditation practice (Grossman & Van Dam, 2011).

A limitation of the majority of effectiveness research and studies included in meta-analyses of MBIs is that they compared MBIs to inactive control conditions which did not control for non-specific elements of interventions (e.g., expectation of benefit, group interaction, facilitator contact), meaning that it is not possible to conclude whether specific aspects of MBIs (i.e., learning mindfulness skills) are responsible for beneficial outcomes. However, recent evidence is emerging which suggests that MBIs may also be effective compared to active control conditions which account for many non-specific elements of interventions; meta-analyses of MBIs compared to active control conditions have found them to be effective for improving depressive symptoms at post-intervention and follow up within a year, with small effect sizes (Strauss, Gu, Kuyken, Cavanagh, & Baer, in preparation), and severity of disorder-specific symptoms in clinical populations at post-intervention and follow up, with small-medium effect sizes (Goldberg et al., 2018).

Mechanisms of Mindfulness-Based Interventions

Given the growing evidence base for MBIs and adaptations for a range of populations, psychological conditions, and outcomes, there is an increasing need to investigate *how* these interventions work; the mechanisms or causal processes which underlie the effects of MBIs on outcomes. An important endeavour is to determine whether as theorised, MBIs work by increasing mindfulness, a key assumption of effectiveness research. In their recent review mapping the evidence base for MBIs to identify ways to increase their public health impact, Dimidjian and Segal (2015) noted that few studies have formally tested mechanisms of MBIs, stating that "underemphasising… precise specification of for whom and how a treatment works risks situating the study of MBI less as science and more as pseudoscience in which mindfulness is seen as a panacea for all problems" (p. 605). Moreover, Kazdin (2007) described several general benefits of identifying mechanisms of psychotherapies, which include the potential to optimise therapeutic change by emphasising active processes of interventions, help identify moderators of treatment and better select suitable patients, facilitate the translation of clinical research to practice, and increase understanding of psychological functioning more broadly, beyond the context of therapy.

Theorised Mechanisms of Mindfulness-Based Interventions

Formal testing of how MBIs work first requires the identification of plausible mechanisms of change which are grounded in theory (Kazdin, 2007). In terms of the theoretical premise of MBSR and MBCT, these MBIs are said to improve outcomes by developing participants' mindfulness skills (Kabat-Zinn, 1982; Segal et al., 2002, 2013). MBCT is also theorised to improve outcomes by increasing disengagement from repetitive negative thinking processes (worry and rumination). Additionally, through consolidating clinical and empirical insights, several overlapping theoretical models and summaries have been proposed which aim to explain how mindfulness and MBIs produce benefits.

In Shapiro, Carlson, Astin, and Freedman's (2006) model, mindfulness, which consists of attention, intention, and attitude, leads to reperceiving, and changes in the following four additional mechanisms: 1) self-regulation, 2) emotional, cognitive, and behavioural flexibility, 3) values clarification, and 4) exposure. They defined

reperceiving as a fundamental shift in perspective whereby one is able to experience mental content (i.e., thoughts) as events in awareness in the moment that they occur and disengage and disidentify from them, viewing present moment experience with greater objectivity and clarity. Another common term for this process is decentering (Safran & Segal, 1990) and metacognitive insight (Teasdale, 1999). Self-regulation is defined as a process of maintaining stability of functioning and adaptability to change. Emotional, cognitive, and behavioural flexibility refers to thinking, feeling, and behaving in ways that are more flexible and adaptive rather than automatic, reactive, and conditioned. Values clarification involves greater recognition of what is meaningful to us so that we can reflectively make decisions that are congruent with our values, interests, and needs. Finally, exposure to difficult mental states leads to desensitisation and understanding that emotions, thoughts, and bodily sensations are not as overwhelming as expected and need not be avoided or feared. Shapiro et al. note that these mechanisms do not represent a linear pathway; instead, each mechanism supports and interacts with others.

Hölzel et al.'s (2011) theoretical framework proposed four mechanisms of action through which mindfulness works: 1) attention regulation, 2) body awareness, 3) emotion regulation, and 4) change in perspective on the self. These four components are said to closely interact to produce enhanced self-regulation. They define attention regulation as the ability to focus and sustain attention whilst disregarding distractions. Body awareness is the ability to recognise subtle bodily sensations. Emotion regulation refers to strategies for regulating emotional responses and Hölzel et al. draw particular attention to two strategies which may be affected by mindfulness practice: reappraisal (changing interpretations of an experience and thus one's emotional response to it) and extinction (exposing oneself to experiences and refraining from avoidance and reactivity). Change in perspective on the self is described as disidentifying with a static sense of self and with the contents of consciousness, and instead perceiving the self as transient. This process is akin to reperceiving (Shapiro et al., 2006) and decentering (Safran & Segal, 1990). Finally, Hölzel et al. note similarities between emotion regulation and change in perspective on the self and compassion for the self.

Vago and Silbersweig (2012) put forward a framework for understanding mindfulness which describes three neurobiological mechanisms through which mindfulness promotes healthy psychological functioning: 1) self-awareness (meta-awareness of self), 2) self-regulation (ability to effectively manage responses and

impulses), and 3) self-transcendence (developing a positive relationship between self and other that transcends self-focus and increases prosocial qualities). These mechanisms are proposed to be supported by six processes: intention, attention and emotion regulation, extinction, prosociality, non-attachment, and decentering.

A further theoretical account of the mechanisms of mindfulness is provided in the Buddhist psychological model (Grabovac, Lau, & Willett, 2011). This model, based on Buddhist psychological theories, highlights five mechanisms underlying the effects of mindfulness practice on wellbeing and symptom reduction: 1) attitudes of acceptance and compassion, 2) attention regulation, 3) ethical practices, 4) decreased attachment and aversion to feelings, and 5) decreased mental proliferation.

Baer (2003) described five suggested mechanisms of mindfulness skills: exposure, cognitive change (e.g., decentering, metacognitive insight), self-management and regulation, relaxation, and acceptance without defence. Similarly, Brown et al. (2007) identified metacognitive insight, exposure, non-attachment (including acceptance and non-avoidance), enhanced mind-body functioning (e.g., greater immunological resistance, relaxation), and integrated functioning (e.g., enhanced executive functioning, self-regulation, autonomy) as processes that may explain the effects of mindfulness.

Taken together, theoretical underpinnings of MBIs and models of mindfulness indicate the involvement of the following mechanisms: increased mindfulness, reperceiving (or decentering, metacognitive insight), self-regulation, attention regulation, emotion regulation, psychological flexibility, values clarification, exposure, body awareness, self-awareness, self-transcendence, compassion, acceptance, non-attachment, relaxation, ethical practices, and mind-body functioning, and reduced rumination, worry, and reactivity. Table 1 in Appendix A presents definitions and known empirical support for these proposed mechanisms.

Methods of Studying Mechanisms

To examine mechanisms, methods of mediation analysis are employed, which test a model in which an independent variable or treatment (X) causes an intervening variable, or mediator (M), which in turn causes a dependent variable or outcome (Y). Over a dozen methods of testing models involving mediator variables have been proposed and these approaches differ in terms of their assumptions, statistical methods

of estimation, theoretical basis, and null hypothesis testing (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002).

The most commonly used approach to investigating the effects of intervening variables is the Baron and Kenny (1986) causal-steps approach, which is based on testing the statistical significance of a series of linear regression models. This approach requires four conditions to be met for there to be mediation: 1) a significant linear relationship between X and M (path *a*), 2) a significant direct relationship between X and Y (path *c*; total effect), 3) a significant relationship between M and Y, after accounting for X (path *b*), and 4) after controlling for M, a relationship between X and Y which is no longer significant (path *c'*; direct effect). The total effect (path *c*) is the sum of the direct effect (path *c'*) and indirect effect (the product of *a* and *b* paths). Although this approach is widely used, it has several limitations. The requirement that the total effect, the relationship between X and Y, has to be significant does not take into account models in which direct and indirect effects have opposite signs and cancel out (MacKinnon et al., 2002). It is also difficult to apply this approach to models which include multiple mediators and determine the effect of each separate mediator (MacKinnon et al., 2002). Additionally, the causal-steps approach attempts to infer the presence of an indirect effect and mediation without directly testing or quantifying the size of the indirect effect (Hayes, 2009; MacKinnon et al., 2002).

Following the causal-steps approach, the most common methods of mediation analysis are the Sobel test, or product of coefficients approach (Sobel, 1982), the distribution of the product approach (MacKinnon et al., 2002), and bootstrapping (Preacher & Hayes, 2004), all of which allow indirect effects to be quantified and significance tests to be conducted. The Sobel test involves dividing the estimate of the indirect effect by its standard error and comparing this value to a standard normal distribution to test for significance. A limitation of the Sobel test is that it assumes normality of the sampling distribution of the indirect effect, which is not typically met (MacKinnon, Lockwood, & Williams, 2004). Neither bootstrapping nor the distribution of the product approach make this assumption, and both have been shown to have greater statistical power and control over Type I error rates compared to the causal-steps method and Sobel test (MacKinnon et al., 2002). Out of these two approaches, bootstrapping has been recommended as the method of choice (Hayes, 2009), due to greater statistical power and more accurate confidence limits being obtained using this method (MacKinnon et al., 2004).

The bootstrapping method involves repeatedly sampling from a dataset and estimating the indirect effect (*ab*) in each resampled data set. It is recommended that this process is repeated at least 5,000 times (Hayes, 2009). This repeated process builds an empirical representation of the sampling distribution of the indirect effect and is used to estimate the size of the indirect effect and generate confidence intervals for the indirect effect. Indirect effects are significant if zero is not contained between the lower and upper boundaries of the bootstrapped confidence intervals.

In addition to the aforementioned methods, Kraemer, Wilson, Fairburn, and Agras (2002) recommended a framework for analysing mediation in RCTs when outcomes are dimensional. Their approach also involves testing linear regression models but differs from the causal-steps method in that to establish mediation, the mediator variable must be correlated with the treatment and have either a main effect on the outcome or an interactive effect with treatment on the outcome.

Drawing conclusions regarding mediation requires more than just conducting methods of mediation analysis; several design requirements should also be attended to. Kazdin (2007) proposed the following seven criteria for establishing mediation: studies should 1) demonstrate a strong association between the intervention and hypothesised mediator and between the mediator and outcome, 2) demonstrate specificity of the mechanism in explaining the association between the intervention and outcome, 3) show consistency by replicating evidence for mediation across studies and samples, 4) use experimental manipulation to strengthen conclusions, ideally manipulating the mediator variable to demonstrate an impact on outcome, 5) establish a timeline to infer causal relation, whereby measurement of mediators temporally precede measurement of outcomes, 6) show a dose-response gradient whereby stronger doses of the mediator is associated with greater change in outcome, and 7) ensure that there is a strong theoretical basis and plausibility for the study of particular variables as mediators (e.g., grounded in theoretical models summarised in the previous section).

Elaborating on the consistency criterion, Kazdin (2007) notes that inconsistency across studies does not necessarily indicate that a variable is not a mediator, as the potential presence of unmeasured moderator variables may affect or mask the mediated relationship. As is implied in the consistency criterion, conclusions regarding mediation is not likely to be reached from a single study and any one study is unlikely to satisfy all design criteria. The case for mediation will require a series of individual mediation studies which together meet all or most of the criteria (Kazdin, 2007) and synthesis of

mediation findings from multiple studies (Cheung & Cheung, 2016; Shadish, 1996). Although there is a growing number of mediation studies of MBIs, narrative and statistical integration of their findings are lacking and this obscures our understanding of the mechanisms of MBIs. This omission is addressed in Chapter 2, which presents a systematic review of mediation studies of MBIs and narrative and statistical syntheses of evidence for identified mechanisms.

Mindfulness-Based Self-Help Interventions

The growing evidence base for the effectiveness of MBIs and investigations of their mechanisms of action have encouraged researchers to begin to explore ways of extending the reach of MBIs to benefit more people. One way of increasing the accessibility of interventions is to offer self-help versions (e.g., learning mindfulness through online programmes, smartphone applications, self-help books, etc.). Self-help interventions are typically briefer, involve reduced administration and service costs, and allow people to work at their own pace and in their preferred environment, all of which help to increase dissemination of intervention content. Mindfulness-based self-help (MBSH) interventions also overcome specific challenges associated with implementing MBIs more widely, such as the shortage of fully trained and accredited MBI therapists and supervision available for MBI therapists (Rycroft-Malone et al., 2017).

A further advantage of MBSH interventions is that they exclude many non-specific elements found in facilitator-led MBIs, such as group support and facilitator contact, making them particularly suited to research examining the specific therapeutic effects of learning mindfulness and associated mechanisms. A common criticism of MBI effectiveness and mechanism research is that lack of comparisons to active control conditions limit conclusions regarding specificity of intervention and mediation effects, and research into MBSH interventions has the potential to inform our understanding of the extent to which, and how, the specific process of learning mindfulness leads to beneficial changes.

Meta-analyses of RCTs of MBSH interventions and acceptance-based self-help interventions have found these interventions to be effective for improving stress, with a medium effect size, and depression, anxiety, wellbeing, and mindfulness with small to

medium effect sizes (Cavanagh, Strauss, Forder, & Jones, 2014; Spijkerman, Pots, & Bohlmeijer, 2016). Like MBIs, these MBSH interventions aim to teach participants mindfulness skills primarily by encouraging them to engage in guided formal and informal mindfulness meditation practices. Such findings are promising and indicate that learning mindfulness using self-help methods has the potential to improve mindfulness, stress, and mental health outcomes. However, on the basis of these findings, it is not yet possible to conclude whether positive changes are entirely and specifically attributable to learning mindfulness skills; findings from these meta-analyses were based largely on studies which compared MBSH interventions to inactive control conditions.

To examine the specificity of the effects of learning mindfulness, not only do we want an MBSH intervention that removes non-specific factors of face-to-face MBIs, we would also need to compare the MBSH intervention to a matched control intervention that controls for additional non-specific elements, such as participant expectations of benefit and taking part in a structured intervention formatted in a similar way and with similar time demands. To our knowledge, only nine studies have compared MBSH interventions to matched control conditions (Carissoli, Villani, & Riva, 2015; Dowd, Hogan, McGuire, Sarma, & Fish, 2015; Howells, Ivtzan, & Eiroa-Orosa, 2016; Jimenez, 2008; Ly et al., 2014; Mongrain, Komeylian, & Barnhart, 2016; Niles et al., 2012; Stankovic, 2015; Wahbeh, Goodrich, & Oken, 2016). Intervention duration in these studies ranged from ten days to eight weeks. Of these studies, four used online MBSH interventions, three audio CDs, and two smartphone applications, and four compared MBSH to psychoeducation, two to relaxation exercises, one to behavioural activation, one to list making, and one to expressive writing. None examined the mechanisms underlying the specific effects of learning mindfulness using methods of mediation analysis.

In addition to there being a need to design a study which has the potential to determine whether the specific process of learning mindfulness leads to improvement in outcomes, such a study should also examine underlying mechanisms, for the aforementioned reasons stated by Kazdin (2007), and to shed light on the integrity of MBSH interventions and differences between MBIs and MBSH interventions; whether these face-to-face and self-help interventions exert their effects in the same way and whether the brief duration of many MBSH interventions (e.g., as little as 10 days) is sufficient for participants to acquire mindfulness skills. This gap is addressed in the

experimental study presented in Chapter 3, which compares the effects of a two-week online MBSH intervention to plausible matched active control and waitlist control conditions and examined theorised mechanisms of change using recommended bootstrapping-based mediation analyses (Hayes, 2009).

Measuring Mindfulness

Valid and reliable ways of measuring mindfulness are necessary for the empirical investigation of the construct, including refining our understanding of mindfulness and how it relates to other constructs, and evaluating the effectiveness and mechanisms of MBIs and MBSH interventions. The vast majority of research in the field have used subjective self-report questionnaire measures of mindfulness, given several advantages of this approach, such as its ability to provide valid and reliable data if the measures used comprehensively capture the construct and are psychometrically robust, and self-report measures being easy to administer and cost-effective. A further reason for the popularity of self-report measures of mindfulness is that to date, no plausible objective measures of mindfulness have been developed (Baer, 2011); no physiological or neurological markers of the general tendency to be mindful have been identified, mindfulness does not seem to be synonymous with the types of abilities assessed in cognitive tests (e.g., sustained attention, working memory), and it does not appear to be easily observable by others and thus not well suited to measures involving direct observation of behaviour. Once promising objective methods of assessing mindfulness are developed, it would be helpful to triangulate them with self-report measures and assess convergence.

Despite the aforementioned advantages of using self-report measures, mindfulness questionnaires have been met with scepticism and it is instructive to briefly consider the main criticisms which are specific to these measures and go beyond well-documented limitations of self-report measures more generally (e.g., social desirability bias when responses are not anonymous or confidential and may have important personal consequences, such as being hired for a job). A concern specific to mindfulness measures is that respondents without any previous meditation experience and who are unaccustomed to attending to their present moment experience may not be able to accurate report on their own capacity to be mindful. A related issue is that

interpretation of mindfulness items may change in the process of practising meditation and differ between people who are unfamiliar with these concepts and those with meditation experience (Grossman, 2008).

Baer (2011) notes that the use of everyday language in self-report measures, accessible to people irrespective of their previous meditation experience, is intended to address these concerns. Such criticisms should not deter researchers from continuing to use, develop, and evaluate self-report measures, given no promising alternative methods, and the potential insights gained from the development and evaluation process, such as determining the presence of these issues in measures (e.g., by investigating differential item functioning and measurement invariance between meditators and non-meditators), refining our understanding of mindfulness, and informing and improving measurement strategies.

Among existing self-report measures of mindfulness, the Five Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006) is one of the most comprehensive and widely used measures of mindfulness (Quaglia, Braun, Freeman, McDaniel, & Brown, 2016). The FFMQ is a 39-item self-report scale measuring the tendency to be mindful in daily life. Its items derived from an exploratory factor analysis (EFA) of the five self-report mindfulness measures available at the time: the Mindfulness Attention Awareness Scale (MAAS; Brown & Ryan, 2003), Freiburg Mindfulness Inventory (FMI; Walach, Buchheld, Buttenmuller, Kleinknecht, & Schmidt, 2006), Cognitive Affective Mindfulness Scale (CAMS; Hayes & Feldman, 2004), Southampton Mindfulness Questionnaire (SMQ; Chadwick, Hember, Symes, Peters, Kuipers, & Dagnan, 2008), and Kentucky Inventory of Mindfulness Skills (KIMS; Baer, Smith, & Allen, 2004). Findings from the EFA indicated that mindfulness comprises five related facets: the capacity to observe internal and external experiences (observing), express in words one's experiences (describing), attend fully to one's present moment activity (acting with awareness), accept rather than judge thoughts and emotions (non-judging), and allow thoughts and emotions to come and go without getting involved or carried away by them (non-reactivity). These facets map well on to the conceptualisation of mindfulness shared by many definitions, as consisting of *what* (observing, describing, acting with awareness) and *how* (non-judging, non-reactivity) (Baer, 2015).

Generally, analyses of the psychometric properties of the FFMQ in samples of regular meditators and non-meditators have demonstrated that this measure has

adequate to excellent convergent validity, internal consistency, interpretability in distinguishing between participant subgroups, incremental validity in predicting mental health symptoms and wellbeing, and sensitivity to change over the course of MBIs (e.g., Baer et al., 2006; Baer et al., 2008; Bohlmeijer, ten Klooster, Fledderus, Veehof, & Baer, 2011; Carmody & Baer, 2008; Christopher, Neuser, Michael, & Baitmangalkar, 2012; Curtiss & Klemanski, 2014a, 2014b). A recent systematic review and meta-analysis of RCTs examining the effects of MBIs on mindfulness when compared to active control conditions, which do not include explicit mindfulness training, has also supported the discriminant validity of the FFMQ (Baer, Gu, Cavanagh, & Strauss, manuscript under review); compared to active control conditions, participants in MBIs showed significantly higher post-intervention mindfulness scores and this was moderated by mindfulness questionnaire used, with significant improvements found for the FFMQ but not for many of the other measures.

Although the psychometric properties of the FFMQ have been supported, findings from recent confirmatory factor analysis (CFA) studies question the validity of its five-factor structure and the inclusion of all five facets in mindfulness research (see Chapter 4 for more information on the following). The five-factor structure which emerged in the development of the FFMQ using EFA (Baer et al., 2006) has been confirmed in meditator samples (Baer et al., 2008; M. J. Williams, Dalgleish, Karl, & Kuyken, 2014), but for non-meditator samples, a four-factor model, with all facets minus 'observing' loading onto an overall mindfulness factor, best fit the data (Baer et al., 2006; Baer et al., 2008; Curtiss & Klemanski, 2014a; M. J. Williams et al., 2014). Inadequate fit of the five-factor model in non-meditators in these studies can be explained by the non-significant relationships found between observing and other facets (non-judging, acting with awareness), which are significant in samples with meditation experience.

Although the factor structure of the FFMQ has been compared in meditators and non-meditators, studies examining the stability of the factor structure before and after an MBI in a single sample is lacking. This is an important omission; the FFMQ is commonly used to assess change before and after MBIs in studies investigating their effectiveness and mechanisms, yet research demonstrating that meditation status results in differential factor structures suggests that such pre-post comparisons may not be valid. This possibility is tested in Chapter 4.

Aforementioned findings, beyond indicating a specific issue with the FFMQ, reinforce the importance of the measure development and evaluation process; empirical testing of scales which conceptualise constructs in a particular way contributes to not only the refinement of the measure, but also to our understanding of the construct itself. For example, observing emerging as part of the FFMQ only for meditator samples suggests that meditation experience may alter people's experience of noticing, by strengthening the relationship between observing experience and other aspects of mindfulness. For people with little or no meditation experience, observing items (e.g., "When I'm walking, I deliberately notice the sensations of my body moving") may reflect neutral attention, or even maladaptive or pathological forms of attention (e.g., anxious monitoring), rather than attention characterised by the *how* qualities (e.g., curious, accepting, and non-judgemental) which are cultivated through mindfulness meditation practice. With meditation experience, people may not only report greater levels of observing, but the way in which they notice may be more consistent with the qualities of mindful attention. Consistent with this explanation, findings show that in non-meditators (student, community, and highly educated samples), observing is the only facet positively correlated with psychological symptoms and in meditators, all facets correlated negatively with psychological symptoms (Baer et al., 2008).

Despite recent findings indicating an issue with the FFMQ factor structure, this can be resolved by excluding the observing facet from comparisons of total scale or subscale scores between meditators and non-meditators (Baer et al., 2006, 2008) and on balance, the FFMQ remains one of the most widely used measures of mindfulness, which not only captures key elements of the construct, but also appears to be the most robust in terms of its psychometric properties. Given these considerations, chapters in this thesis use the FFMQ to measure mindfulness.

Compassion in Context

"A human being is part of a whole, called by us the "Universe," a part limited in time and space. He experiences himself, his thoughts and feelings, as something separated from the rest — a kind of optical delusion of his consciousness. This delusion is a kind of prison for us, restricting us to our personal desires and to affection for a few persons nearest us. Our task must be to free ourselves from this prison by widening our circles of compassion to embrace all living creatures and the whole of nature in its beauty. Nobody is able to achieve this completely, but the striving for such achievement is in itself a part of the liberation and a foundation for inner security."
– Albert Einstein

Compassion is upheld as a core human virtue in major contemplative and religious traditions (Dahlsgaard, Peterson, & Seligman, 2005). In Buddhist traditions, the cultivation of compassion is inextricably linked to the development of mindfulness and aspired to as the central goal of many contemplative practices. Mindfulness training is viewed as a necessary foundation for building compassion for the self, and ultimately compassion for all beings (Thurman, 1997). MBIs such as MBSR and MBCT may implicitly develop compassion through teachers' embodiment of this quality, learning to decenter from mental content, group enquiry sessions, and encouragement to approach difficult experiences with openness, acceptance, non-judgement, and curiosity (Feldman & Kuyken, 2011). Development of mindful awareness and insight may also dissolve our sense of separation from other people so that we increasingly view others as fundamentally like ourselves, strengthening our compassion for them. In support of mindfulness and compassion being intertwined, studies have found strong correlations between self-reported levels of mindfulness and self- and other-compassion (e.g., Gu et al., under review). Research has also demonstrated increases in compassion and related outcomes following participation in MBIs (e.g., Birnie, Speca, & Carlson, 2010; Kuyken et al., 2010; Lim, Condon, & DeSteno, 2015; Robins, Keng, Ekblad, & Brantley, 2012; Shapiro, Astin, Bishop, & Cordova, 2005).

In recent decades, there has been a rapidly growing societal and scientific interest in compassion and recognition of its importance. In particular, there has been a focus on creating a culture of compassion in the healthcare sector, with compassion being one of the six fundamental values in the UK National Health Service (NHS) constitution (Department of Health, 2013) and central to the American Medical

Association's (AMA) Principles of Medical Ethics (AMA, 2001). Treating patients compassionately is argued to have numerous advantages, including improving clinical outcomes, patient satisfaction with services, and the quality of information gathered from patients (Epstein et al., 2005; Redelmeier, Molin, & Tibshirani, 1995; Sanghavi, 2006). The importance of treating oneself compassionately has also been acknowledged in this domain, with large scale programmes aimed at improving the health and wellbeing of healthcare staff (NHS England, 2017). The important role of compassion in other sectors has also been recognised, with calls and initiatives to integrate teaching of compassion in schools (e.g., Compassion in Education Foundation, 2016) and acknowledgement of compassion as being core to the ethical framework guiding justice systems (e.g., Norko, 2005).

Research has linked enhanced compassion to a number of benefits, such as greater wellbeing (Davidson & Schuyler, 2015), increased happiness and self-esteem (Mongrain, Chin, & Shapira, 2011), increased social connectedness and reduced loneliness (Crocker & Canevello, 2008), and reduced depressive symptoms and negative affect (López, Sanderman, Ranchor, & Schroevers, 2018).

Compassion-Based Interventions

In light of widespread recognition of the importance of compassion and benefits associated with compassion, a number of psychological interventions have been developed which aim to specifically enhance people's capacity for compassion for themselves and other people. A recent review of compassion-based interventions (CBIs) identified five empirically supported programmes that focus on the cultivation of compassion (Kirby, 2017): 1) Compassion Focused Therapy (CFT; Gilbert, 2014), 2) Mindful Self-Compassion (MSC; Neff & Germer, 2013), 3) Compassion Cultivation Training (CCT; Jinpa, 2010), 4) Cognitively-Based Compassion Training (CBCT; Pace et al., 2009), and 5) Cultivating Emotional Balance (CEB; Kemeny et al., 2012). Although these interventions differ in terms of how they define compassion (Kirby, 2017), and therefore the precise capacities they aim to cultivate, they share a number of features. Theoretically, all of these CBIs are underpinned by Tibetan Buddhist teachings but have been designed to be secular in approach. As such, all include training in mindfulness as a foundation for the cultivation of compassion. CBIs are

formatted in a similar way to MBIs; they are typically group-based and consist of eight or nine weekly sessions, with each session ranging from 2 to 2.5 hours. Interventions are multimodal, and mindfulness and compassion skills are developed through a combination of meditation practices, behavioural exercises, psychoeducation, and homework exercises. Compassion meditation practices typically involve contemplating feelings of compassion for the self or close others, and gradually extending these feelings to less close others, adversaries, and all beings. An assumption of these programmes is that a dispositional trait-like tendency to be compassionate towards others and/or oneself can be developed through repeated practice of skills that cultivate compassionate states.

A recent meta-analysis of 21 RCTs found that compared to both active and inactive control conditions, CBIs are effective in improving mindfulness, compassion for the self, depression, anxiety, wellbeing, and psychological distress, with moderate between-group effect sizes (Kirby et al., 2017). Kirby et al.'s meta-analysis also found CBIs to be effective for improving compassion for others compared to waitlist conditions, but this finding was based on just four studies, three of which used direct self-report measures of the construct. This is a significant limitation; without directly measuring compassion, we cannot ascertain whether many of these CBIs, purported to cultivate compassion for the self and/or others, are effective and whether beneficial effects on other outcomes are the result of improvements in the theorised mechanism (i.e., compassion) rather than due to placebo effects or changes in other constructs. This also impedes future refinement of these interventions; for example, we are unable to determine adequate dose of intervention components (e.g., how much teaching of other-compassion versus self-compassion is needed) or assess the learning trajectory (e.g., whether initial training in compassion for the self is supportive of the development of compassion for others).

The paucity of RCTs of CBIs which included a measure of compassion is indicative of issues around conceptualising and measuring compassion in the field. These issues are also reflected in the various ways in which interventions have defined, and therefore aim to cultivate, compassion. The following two sections give an overview of existing conceptualisations and measures of compassion, highlight key omissions in these areas, and call for greater clarity and consensus on how to define and measure compassion constructs as necessary initial steps for progressing research in this burgeoning field.

Conceptualising Compassion

Despite the recognised importance of compassion and increasing research and societal interest, a number of definitions exist and there is lack of consensus on the key features which comprise compassion. In order for this field to develop, there is a need to consolidate existing conceptualisations and need for common language and understanding with which to refer to compassion constructs.

In Buddhist traditions, compassion is regarded as a response to suffering, as "the heart that trembles in the face of suffering" (Feldman & Kuyken, 2011, p. 144). The Pali term *karuna* is viewed as the most direct translation of compassion (Rahula, 1959) and involves a desire to prevent suffering from occurring to the self and others (Tirch, 2010). Thus, within Buddhism, compassion is viewed as more than just an emotional response to suffering; it is a response that is embedded in an ethical framework dedicated to freeing the self and others from suffering. The Dalai Lama's (1995) definition of compassion is consistent with this approach; compassion entails a sensitivity to the suffering of self and others with a deep wish and commitment to relieve the suffering. Inherent in the Buddhist conceptualisation of compassion is that it is a response experienced specifically in relation to suffering and the lack of distinction between compassion for the self and others; the capacity to meet suffering with compassion is the same irrespective of target.

By contrast, the Oxford English Dictionary defines compassion solely in terms of an emotional response to the suffering of others ("sympathetic pity and concern for the sufferings or misfortunes of others") and notes that the word 'compassion' derives from the Latin *compati*, meaning "to suffer with". However, among emotion researchers, only 20% agree that compassion is an emotion (Ekman, 2016). Additionally, there is a need to distinguish compassion from related states; to move away from using the term compassion interchangeably with terms such as sympathy, pity, and empathy, and avoid using these related terms in definitions of compassion. Findings from Goetz, Keltner, and Simon-Thomas' (2010) review indicate that compassion is a distinct construct, which can be distinguished from related states in a number of ways; in terms of the appraisal processes, affective experience, physiological responses, and patterns of behaviour involved.

In psychological literature, consistent with the Buddhist conceptualisation, compassion is generally seen as multidimensional and consisting of cognitive, affective, and behavioural elements. Kanov et al. (2004) identify three key elements of compassion: noticing, feeling, and responding. Noticing involves cognitively recognising a person's suffering or may first be experienced through an unconscious physical or affective reaction to distress which produces an awareness of suffering. Feeling is defined as emotionally connecting to that suffering and experiencing empathic concern through imagining or feeling the condition of the person suffering. Responding refers to the desire to act or acting to ease the person's suffering. Similarly, Geshe Thupten Jinpa, who developed the CCT intervention, conceptualises compassion as comprising four components: 1) an awareness of suffering, 2) being emotionally moved by suffering, 3) a wish and intention to ease suffering, and 4) a responsiveness or readiness to act to relieve suffering (Jazaieri et al., 2013).

Gilbert (2009a), who developed CFT, defines compassion as: "A deep awareness of the suffering of another coupled with the wish to relieve it" (p. 13) and views compassion as consisting of six attributes: sensitivity, sympathy, empathy, care for wellbeing, distress tolerance, and non-judgement (Gilbert, 2010). Sensitivity involves being responsive to another's distress and being able to perceive their needs. Empathy refers to being able to understand another person's point of view or our own thoughts and feelings. Sympathy involves being emotionally moved by the distress of another. Care for wellbeing is the motivation to act or facilitating action to help alleviate suffering. Distress tolerance involves the ability to stay with difficult emotions in oneself when faced with someone else's suffering without avoiding or becoming overwhelmed by them. Lastly, non-judgement means remaining accepting of and tolerant towards another person rather than criticising, shaming, or rejecting. Distress tolerance and non-judgement are important aspects of compassion because if we feel overwhelmed by or are judgemental about a person's suffering, we may experience personal distress and become self-focused, which may prevent a compassionate response and lead instead to actions to reduce our distress, such as avoiding the person suffering or reducing our awareness of their pain (Gilbert, 2010).

Wispe (1991) also conceptualises compassion as not only recognising, feeling moved by, and wanting to alleviate suffering, but also involving the ability to be non-judgmental towards others and to tolerate our own distress when confronted with another's suffering. Additionally, distress tolerance and non-judgement are inherent in

Buddhist conceptualisations. For example, the Dalai Lama (2002b) stated that: “for a practitioner of love and compassion, an enemy is one of the most important teachers. Without an enemy you cannot practice tolerance, and without tolerance you cannot build a sound basis of compassion” (p. 75).

Finally, Neff (2003a) developed a model of self-compassion which consists of three components: kindness (being kind and non-judgmental towards the self rather than self-critical), mindfulness (holding difficult experiences in mindful awareness rather than over-identifying with them), and common humanity (viewing suffering as part of the human condition rather than as isolating). Pommier (2010) later developed a model of compassion for others based on Neff’s model of self-compassion which consists of the same three components: being kind and understanding towards others who are suffering rather than critical or indifferent (kindness), noticing and remaining open to another’s suffering without over-identifying with their distress (mindfulness), and understanding that suffering is a shared human experience (common humanity).

The emphasis in these models on suffering being a common human experience is also present in Buddhist teachings of compassion. The Buddhist view of the self emphasises an interdependent nature of existence (Tirch, 2010) and Wang (2005) defines compassion as “the feeling that arises from the realisation of the deeper reality that we are all connected, we are all one” (p. 104). Similarly, Feldman and Kuyken (2011) describe compassion as: “an orientation of mind that recognises pain and the universality of pain in human experience and the capacity to meet that pain with kindness, empathy, equanimity and patience” (p. 145).

Taken together, it is clear from this overview of a subset of key definitions that there are differences in how compassion has been defined in the field. Given these varying definitions, it is not surprising that several different CBIs have been developed, each aiming to cultivate a different set of capacities which are regarded as comprising compassion. In the context of increasing and widespread interest in compassion and how it can be enhanced, there is a need to consolidate the range of conceptualisations of compassion into one comprehensive, operational definition. Such a unifying definition is necessary for progressing key research in this area, for instance developing valid and reliable measures, using these measures to evaluate CBIs, and refining and optimising the effectiveness of CBIs. Chapter 5 addresses this gap in compassion research, by reviewing and consolidating the range of conceptualisations of compassion into one multifaceted definition.

Measuring Compassion

As with there being a number of CBIs which reflect the range of ways in which compassion has been conceptualised, numerous measures of compassion have also been developed which vary in the key features of compassion they capture (e.g., acting to alleviate suffering, recognising suffering, feeling moved by suffering).

Behavioural measures of compassion include prosocial games and experimental situations which assess whether an individual engages in helping behaviour. An example of a prosocial game is the Zurich Prosocial Game (ZPG; Leiberg, Klimecki, & Singer, 2011) which assesses prosocial behaviour and the degree to which this is attenuated by factors such as reciprocity of help, helping cost, and the presence of distress cues. In short, participants are instructed to navigate a virtual character through a maze and reach a treasure in a limited amount of time. They play at the same time as a pre-programmed co-player, described as being from another research institute in Europe, and can view the path of the other player. The two players do not share the same paths in the maze and do not compete for the same treasure. During the game, gates fall on the paths that can block the participant and co-player and the participant can choose to use keys they possess to open gates for the co-player, thus helping them in their game. Leiberg et al. found that this helping behaviour increased when it is reciprocated (co-players previously used their keys to help the participant), there are distress cues (cues conveyed by the co-player's character, e.g., crying sounds heard over headphones), and the cost of helping is low (participants do not need the keys they donate). Helping behaviour in the ZPG was also found to significantly increase in participants who took part in a one-day compassion training course, but not in those assigned to an active control condition (one-day memory training) (Leiberg et al., 2011).

Lim et al. (2015) devised an experimental task assessing whether an individual engages in helping behaviour. Their task assessed whether participants who arrive at a waiting area and occupy one of three seats (the other two taken up by confederates) offer their seat to a third confederate, who, arriving after the participant is seated, is in crutches and displaying discomfort. Lim et al. found that participants who completed

three weeks of mindfulness training were significantly more likely to give up their seats compared to those in the active control condition who took part in cognitive training.

In terms of other objective markers of compassion, research has found structural changes in frontoinsular brain regions following compassion training (Valk et al., 2017) and implicated increased heart rate variability (HRV), an index of parasympathetic nervous system activity, as a physiological indicator of enhanced prosocial behaviour and compassion (Porges, 2009).

Although it is helpful to explore and develop objective and behavioural measures of compassion, such measures are limited in a number of ways. In the case of behavioural measures, they provide only a partial picture of compassion by solely assessing helpful responding. Many definitions of compassion also include an emotional element (e.g., emotionally connecting with suffering), cognitive element (e.g., recognising and understanding the universality of suffering), and motivational element (e.g., motivation to act to alleviate suffering) which are collectively regarded as driving helping behaviour and it is possible that the helpful behaviour observed in behavioural tasks are the result of alternative motivations (e.g., acting in ways that are socially appropriate or desirable). Objective neurological and physiological markers of compassion also reveal nothing about a person's qualitative experience of compassion. Additionally, both behavioural and objective measures are limited in terms of specificity; it is unclear at present whether these tools are measuring compassion or other related states which may not comprise emotional, cognitive, and motivational factors (e.g., empathy, sympathy, kindness, pity).

Given the difficulties of assessing a multidimensional construct such as compassion using solely behavioural and objective methods, research in this field has most commonly measured compassion using self-report questionnaires. Although these are not without their limitations (e.g., response biases), self-report scales can be valid and reliable if they have been developed to comprehensively capture compassion, items have been generated in line with good practice guidelines, care has been taken to ensure items do not capture related constructs, and scales have been demonstrated to be psychometrically robust.

A number of self-report measures of compassion exist, and these differ in how they conceptualise compassion and the elements of compassion captured by items. For example, the Compassionate Love Scale (Sprecher & Fehr, 2005) measures compassion for close others and for humanity at large and includes items which were

generated based on a review of the literature on love and altruism. This scale assesses emotional connection, common humanity, and motivation to help, but does not capture people's recognition or awareness of suffering. It additionally includes the term 'compassion' in the items and not all items refer specifically to suffering.

The Compassionate Engagement and Action Scales, recently developed by Gilbert et al. (2017), measure self-compassion, compassion for others, and compassion from others. Items from these scales capture 'compassionate action' plus the six attributes from Gilbert's (2010) definition which collectively comprise 'compassionate engagement': sensitivity, sympathy, empathy, care for wellbeing, distress tolerance, and non-judgement. Although this scale captures a number of key elements of compassion, it does not include common humanity, and there are several methodological limitations, such as items being double-barrelled, the generation of just one item per attribute, and EFA being conducted within each of the two factors (engagement and action) rather than across all items to determine the underlying conceptual structure.

Another measure of self-compassion is the widely-used self-compassion scale (SCS; Neff, 2003b). Development of items for this scale was based on Neff's three-component conceptualisation of self-compassion as comprising self-kindness, mindfulness, and common humanity. However, a number of factor analytic studies have recently indicated issues with the conceptual structure of the SCS (e.g., Costa, Marôco, Pinto-Gouveia, Ferreira, & Castilho, 2015; López et al., 2015; M. J. Williams et al., 2014).

Taken together, various self-report measures of compassion exist but there is a lack of clarity in terms of their quality and which to use. Continued use of measures which are not valid and/or reliable can undermine research findings and be counterproductive for the field. There is therefore a need to review scales both in terms of whether their items adequately capture all facets of compassion (content validity), in line with a comprehensive definition of compassion, and their psychometric properties. If current measures are found to be lacking in terms of whether they accurately and comprehensively capture compassion and their psychometric properties, then this would necessitate the development and validation of new measures of compassion. Issues around measurement clarity are addressed in Chapter 5, which in addition to reviewing and consolidating definitions of compassion, also includes a systematic

review and narrative synthesis of questionnaire measures of compassion for the self and others.

The systematic review in Chapter 5 identified significant limitations with existing compassion measures both in terms of their content validity and psychometric properties. This calls for the development of psychometrically robust measures of compassion for the self and others, which are based on a comprehensive, theoretically and empirically informed definition. These important omissions are addressed in the final two chapters of this thesis; Chapter 6 empirically tests the definition of compassion derived from Chapter 5, and Chapter 7 uses the theoretically and empirically informed definition as a basis for the development and validation of new self-report measures of self- and other-compassion.

The Current Thesis

This thesis aims to address several important gaps in mindfulness and compassion research which hinder progress in these fields and limit the public health impact of MBIs and CBIs. In the case of mindfulness, although a growing number of studies have examined the mechanisms of MBIs, narrative and statistical synthesis of their findings are lacking, and this is needed to assess the consistency of findings and make robust conclusions about how MBIs work. Establishing whether a variable is a mechanism also requires attending to the design of individual studies. A benefit of evaluating MBSH interventions is that they exclude many non-specific elements found in MBIs, lending themselves well to research examining the specific therapeutic effects and associated mechanisms of learning mindfulness. However, few studies have evaluated the effectiveness MBSH interventions compared to matched control conditions, which control for additional non-specific factors, and none have used this design to investigate mechanisms. This approach has the potential to greatly benefit the study of mechanisms, by shedding light on how the specific process of learning mindfulness leads to improvement in outcomes. Finally, research in this field requires valid and reliable measures of mindfulness. In the context of mechanism and effectiveness research, these measures need to maintain a stable factor structure when they are used to assess change before and after MBIs. However, recent research suggests that this may not be the case for a widely-used and otherwise

psychometrically robust measure of mindfulness, but this possibility requires empirical testing.

Despite increasing research and societal interest in compassion for others and compassion for the self, these constructs lack basic definitional and measurement clarity. There is a need to consolidate the range of conceptualisations of compassion into one comprehensive, operational definition and to review existing self-report scales in terms of their content validity and psychometric properties. There is some indication that existing measures may lack content validity and/or have poor psychometric properties and if these issues are consistent across existing measures, this would warrant the development and validation of new measures of compassion.

Overview of Chapters

The following chapters of this thesis collectively address the aforementioned gaps in the literature on mindfulness and compassion. Chapters 2, 3, and 4 address the omissions in mindfulness research by examining the mechanisms of MBIs and measurement of mindfulness. Chapter 2 presents a systematic review and meta-analysis of studies which formally tested mechanisms of MBIs using mediation analysis. This chapter evaluates the strength and consistency of evidence for each mechanism identified in the systematic review. For identified mechanisms with sufficient evidence, two-stage meta-analytic structural equation modeling is used to combine findings from mediation analyses, to examine whether these mechanisms significantly mediate the effect of MBIs on mental health outcomes. Chapter 3 examines the specific effects and mechanisms of learning mindfulness on stress in a sample of University students and staff, by comparing a two-week online self-help MBI with a matched control condition and controlling for non-specific effects. The mechanisms tested were mindfulness, self-compassion, and worry, and these were selected because they have been identified as theoretically and empirically supported mechanisms of MBIs. Chapter 4 uses CFA to assess the stability of the factor structure of the widely used FFMQ before and after MBCT in a sample of adults with recurrent depression in remission.

Chapters 5, 6, and 7 address the omissions in compassion research around the lack of definitional and measurement clarity. Chapter 5 presents a review of definitions of compassion and a systematic review and evaluation of nine existing compassion questionnaire measures. Following the review of conceptualisations of compassion, this

chapter proposes a five-element definition of compassion, which captures the key features of compassion and can be applied to both self- and other-compassion. The subsequent systematic review of compassion measures identifies limitations with existing compassion measures both in terms of their content validity and psychometric properties. Findings from this chapter call for the empirical testing of the five-element definition and if supported, the development and validation of new measures of compassion for the self and others based on this theoretically and empirically supported definition. Chapter 6 empirically investigates the factor structure of the five-element definition using self-report items. Both EFA and CFA are used to first identify the underlying structure of compassion and then to validate the identified factor structure. Following support for the five-element conceptualisation of compassion, Chapter 7 uses this theoretically and empirically informed definition as a basis for the development and validation of new self-report measures of self- and other-compassion. The resulting measures are developed and validated across four stages: 1) item generation and review through expert and non-expert consultation, 2) item reduction, 3) CFA in a sample of healthcare staff to validate the factor structure, and 4) CFA in a sample of University students to cross-validate the factor structure. Finally, Chapter 8 summarises all of the findings, evaluates the strengths and limitations of this programme of research, and discusses future directions and implications.

Chapter 2:
How Do Mindfulness-Based Cognitive Therapy and Mindfulness-Based Stress Reduction Improve Mental Health and Wellbeing? A Systematic Review and Meta-Analysis of Mediation Studies

Chapter 2 is published in *Clinical Psychology Review* as:

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical psychology review, 37*, 1-12.

Author Contributions

Design:	JG, CS, and KC
Data Collection:	JG
Data Analysis:	JG and RB
Initial Draft:	JG
Reviewing and Editing:	All authors

Abstract

Given the extensive evidence base for the efficacy of mindfulness-based stress reduction (MBSR) and mindfulness-based cognitive therapy (MBCT), researchers have started to explore the mechanisms underlying their therapeutic effects on psychological outcomes, using methods of mediation analysis. No known studies have systematically reviewed and statistically integrated mediation studies in this field. The present study aimed to systematically review mediation studies in the literature on mindfulness-based interventions (MBIs), to identify potential psychological mechanisms underlying MBCT and MBSR's effects on psychological functioning and wellbeing, and to evaluate the strength and consistency of evidence for each mechanism. For the identified mechanisms with sufficient evidence, quantitative synthesis using two-stage meta-analytic structural equation modeling (TSSEM) was used to examine whether these mechanisms mediate the impact of MBIs on clinical outcomes. This review identified moderate and consistent evidence for mindfulness, rumination, and worry, and preliminary but insufficient evidence for cognitive and emotional reactivity, self-compassion, and psychological flexibility as mechanisms underlying MBIs. TSSEM demonstrated evidence for mindfulness, rumination and worry as significant mediators of the effects of MBIs on mental health outcomes. Most reviewed mediation studies have several key methodological shortcomings which preclude robust conclusions regarding mediation. However, they provide important groundwork on which future studies could build.

Keywords: mindfulness, mechanisms, mediation, structural equation modeling, meta-analysis, mental health

Introduction

Although there are diverse definitions of mindfulness, it is commonly and operationally defined as the quality of consciousness or awareness that arises through intentionally attending to present moment experience in a non-judgemental and accepting way (Kabat-Zinn, 1994). Mindfulness originates from Eastern traditions and its recent popularity in Western psychology is largely due to the development and widespread application of standardised mindfulness-based interventions (MBIs), which integrate the essence of traditional mindfulness practice with contemporary psychological practice, in order to improve psychological functioning and wellbeing. The two most extensively employed and evaluated MBIs are mindfulness-based stress reduction (MBSR; Kabat-Zinn, 1982) and mindfulness-based cognitive therapy (MBCT; Segal, Williams, & Teasdale, 2002; 2013), both of which are eight-week group-based therapies which teach mindfulness skills through a range of formal and informal mindfulness practices. These include mindfulness of breath, thoughts, bodily sensations, sounds, and everyday activities. A growing body of robust evidence from randomised controlled trials (RCTs) has demonstrated that MBIs are effective in improving a range of clinical and non-clinical psychological outcomes in comparison to control conditions, including anxiety (Green & Bieling, 2012; Hofmann, Sawyer, Witt, & Oh, 2010), risk of relapse for depression (Kuyken et al., 2008; Teasdale et al., 2000), current depressive symptoms (Strauss, Cavanagh, Oliver, & Pettman, 2014), stress (Chiesa & Serretti, 2009), chronic pain (Grossman, Tiefenthaler-Gilmer, Raysz, & Kesper, 2007), quality of life (Godfrin & van Heeringen, 2010; Kuyken et al., 2008), psychological symptoms in patients with cancer (Ledesma & Kumano, 2009), and retrieval of specific autobiographical memories (Williams, Teasdale, Segal, & Soulsby, 2000), a reliable cognitive marker of depression (e.g., Brittlebank, Scott, Williams, & Ferrier, 1993).

Other notable interventions which involve mindfulness principles alongside other components include acceptance and commitment therapy (ACT; Hayes & Wilson, 1994) and dialectical behavioural therapy (DBT; Linehan, 1993). A consensus has not yet been reached regarding the similarity between MBSR and MBCT and these alternative interventions (Chiesa & Malinowski, 2011). However, it is clear that alternative interventions differ considerably from MBSR and MBCT and each other in the duration and frequency of mindfulness practice involved, and the inclusion of

mindfulness psychoeducation and non-mindfulness therapeutic ingredients. Therefore, the current review will focus solely on MBCT and MBSR to ensure consistency across studies and will use the term MBI to refer to these two interventions. These interventions have much in common in their core elements, have published therapy protocols that are adhered to in research trials, and have generated a large body of evidence.

Hypothesised Mechanisms Underlying Mindfulness-Based Interventions

Compared to the extensive evidence base for the effectiveness of MBSR and MBCT, relatively few studies have tested *how* these MBIs work, or the mechanisms of action which causally connect changes that occur during MBIs with psychological outcomes. Kazdin (2007) emphasises several clinically relevant reasons why establishing the mechanisms of psychotherapies is crucial. These include being able to optimise therapeutic effects through enhancing active components of interventions, distinguishing between the specific and broader, non-specific effects of treatment, facilitating the identification of treatment moderators and matching of therapies to individuals, and informing theory development and interpretation of results. Moreover, Brown, Ryan, and Creswell (2007) state that developing a deeper understanding of mindfulness, including how it works, is a worthy venture as it has the potential to reveal insights into the age-old mystery of the nature of consciousness.

Studies which have examined the mechanisms of MBIs have typically based their investigation on the theoretical underpinnings of MBSR and MBCT. A theoretical premise of MBCT and MBSR is that the development of mindfulness skills leads to non-judgemental and non-reactive acceptance of all experience, which in turn results in positive psychological outcomes (Kabat-Zinn, 1982; Segal et al., 2002, 2013). Additionally, MBCT, originally developed with the intention of treating recurrent depression, is theorised to decrease depressive recurrence by enhancing awareness of and disengagement from repetitive negative thinking about one's depressive symptoms (Segal et al., 2002, 2013). A further cognitive process proposed and found to be implicated in the relationship between MBCT and depression (Williams et al., 2000) and closely linked to ruminative analytic thinking (Raes et al., 2006; Watkins & Teasdale, 2001, 2004) is autobiographical memory specificity (AMS), which refers to the ability to retrieve memories of specific personal events that happened at particular

times and locations. MBCT, which encourages participants to notice specific details of their environment and accept all experiences without judging or avoiding them, is likely to improve AMS through reducing overgeneric encoding of situations and suppression of unpleasant memory retrieval (Williams et al., 2000). Therefore, we might expect MBSR and MBCT to improve psychological outcomes through increasing levels of mindfulness and non-judgemental acceptance, and through decreasing negative reactivity, repetitive negative thinking, and overgeneral autobiographical memory retrieval.

Additionally, through integrating knowledge from empirical studies of MBIs, several researchers have developed theoretical models or summaries that include a wider range of potential mechanisms of mindfulness and MBIs. For instance, Shapiro, Carlson, Astin, and Freedman's (2006) model proposes that mindfulness, composed of attention, intention and attitude, leads to reperceiving and changes in the following four mechanisms: 1) self-regulation, 2) emotional, cognitive and behavioural flexibility, 3) values clarification, and 4) exposure. Hölzel and colleagues' (2011) theoretical review integrated neuroscientific findings with self-report and experimental data to propose four mechanisms through which mindfulness works: 1) attention regulation, 2) body awareness, 3) emotion regulation, and 4) change in perspective on a 'static' self. They also draw similarities between emotion regulation and change in perspective on the self with self-compassion, a construct conceptualised by Neff (2003b) as being is closely linked to mindfulness and consisting of three components: self-kindness in the face of suffering, seeing one's experience as part of larger human experience, and 'mindfulness', defined as "holding one's painful thoughts and feelings in balanced awareness rather than over-identifying with them" (p. 224). Vago and Silbersweig's (2012) framework and neurobiological model describes three mechanisms through which mindfulness promotes positive mental health and reduces biases related to self-processing: 1) self-awareness, 2) self-regulation, and 3) self-transcendence. Brown and colleagues (2007) also describe several processes underlying the therapeutic effects of mindfulness, including insight, exposure, nonattachment, enhanced mind-body functioning, and integrated functioning. Similarly, Baer (2003) identified exposure, cognitive change, self-management, relaxation, and acceptance as key mechanisms. A further model of the mechanisms of mindfulness is the Buddhist psychological model (Grabovac, Lau, & Willett, 2011). This model, based on Buddhist texts, proposes acceptance/compassion, attention regulation, ethical practices, nonattachment and

nonaversion, and decreased mental proliferation as mechanisms underlying the effects of mindfulness practice on clinical symptom reduction and wellbeing. The proposed mechanisms of mindfulness and MBIs in these theoretical models are presented in Table 1.

Taken together, based on the theoretical underpinnings of MBIs and models of how they work, possible mechanisms connecting MBSR and MBCT with their beneficial effects include improvements in a number of variables including mindfulness, repetitive negative thinking, AMS, reperceiving, reactivity, nonattachment, nonaversion, self-awareness, self-regulation, self-transcendence, psychological flexibility, clarification of inner values, exposure, attentional control and regulation, body awareness, mind-body and integrated functioning, emotion regulation, self-compassion, compassion, insight, acceptance, relaxation, and ethical practices.

Methods of Studying the Mechanisms Underlying Mindfulness-Based Interventions

To understand how change occurs during interventions, conducting mediation analyses to study the indirect influence of a treatment (X) on an outcome (Y) through a mediator (M), or intervening variable, is an essential first step (Kazdin, 2007). There are over a dozen methods of mediation analysis, most of them testing the statistical significance of a sequence of linear regression models (Baron & Kenny, 1986), or using structural equation modeling (SEM), which allows simultaneous examination of direct and indirect relationships among constructs represented by multiple items (Kline, 2011). Researchers have advocated the use of SEM techniques for assessing mediation (e.g., Preacher & Hayes, 2004) and empirically demonstrated their superiority over regression procedures (Iacobucci, Saldanha, & Deng, 2007). Nevertheless, the most popular method of mediation analysis, the Baron and Kenny (1986) causal-steps test, uses a regression framework. Under this approach, mediation is said to occur if four conditions are fulfilled through conducting a series of regression analyses: 1) there is a significant linear relationship between X and M (path *a*), 2) there is a relationship between X and Y (path *c*), 3) there is a relationship between M and Y, after accounting for X (path *b*), and 4) after controlling for M, the relationship between X and Y decreases in size relative to the size of the regression coefficient for path *c* (path *c'*). Full mediation refers to when the regression coefficient for path *c'* is not significantly

different from zero and partial mediation is said to occur if this coefficient has decreased in size relative to the coefficient for path *c* but remains significant (Baron & Kenny, 1986).

Following the causal-steps approach, the most common methods of mediation analysis and testing the significance of indirect effects, based on regression or SEM frameworks, are the Sobel test, or product of coefficients approach (Sobel, 1982), the distribution of the product approach (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002), and the bootstrapping method (Preacher & Hayes, 2004). Of the most popular methods, MacKinnon and colleagues recommended bootstrapping or the distribution of the product approach over the Sobel test or causal steps approach, based on the former having greater statistical power and the most accurate Type I error rates (see MacKinnon et al. (2002) for a detailed evaluation of different methods of mediation analysis). Additionally, Kraemer, Wilson, Fairburn, and Agras (2002) recommended an analytic framework for testing mediation in RCTs when outcomes are dimensional. The framework shares the regression basis with the causal steps approach but differs in that a variable is demonstrated to be a mediator when it is correlated with the treatment and has either a main effect on the outcome or an interactive effect with treatment on the outcome.

In addition to there being a wide range of methods for testing mediational hypotheses, some recommended over others, drawing inferences about a mediator also involves several design requirements. Kazdin (2007) postulated that ideally, to establish mediation, there should be a strong theoretical basis for the study of certain variables as mediators, studies should compare the intervention to a control group to make a strong case for specificity of effects to treatment, measure change in mediators during the intervention and before outcomes, and show consistency by replicating evidence for mediation across studies. Inconsistency however, does not necessarily mean that a mediator is not involved, due to the potential presence of unmeasured moderator variables affecting the mediated relationship. Lastly, studies should analyse data from only participants who receive a sufficient dose of the intervention (Kazdin, 2007), commonly defined in the MBI literature (e.g., Kuyken et al., 2008; Teasdale et al., 2000) as participation in more than four out of eight weekly sessions. Not adhering to such methodological requirements results in conclusions about mediation which are premature and tentative at best. Only after studies meeting all or most of these criteria

have consistently demonstrated that a particular variable accounts for the effects of an MBI on an outcome can this variable be established as a mechanism.

The Current Study

Despite systematic reviews long being advocated as a method for investigating the mediators attributed to interventions (Shadish, 1996), as interpretation of what explains a phenomenon cannot emerge from a single study (Kazdin, 2007), there are no known systematic reviews of the mechanisms of MBIs reviewing only mediation studies. Researchers have also advocated integrating meta-analytic techniques with SEM to statistically evaluate and synthesise evidence for mediation across multiple studies (e.g., Cheung, 2008; Shadish, 1996). Cheung and Chan (2005) proposed a two-stage SEM (TSSEM) method to unify meta-analysis and SEM in a fixed or random-effects model, by first, synthesising correlation matrices between X, M, and Y across studies and second, fitting structural mediational models on the pooled correlation matrix.

The present study's aims are threefold. First, it aims to systematically review mediation studies in the literature on MBIs and identify the mechanisms underlying MBCT's and MBSR's effects on psychological outcomes. Second, it aims to provide narrative summaries of the strength and consistency of evidence for each identified mechanism. Third, it aims to conduct separate TSSEM analyses for each identified mechanism of the effects of MBIs on mental health outcomes which have been supported by a substantial pool of evidence. There is currently no consensus in the multivariate meta-analysis literature on the minimum number of studies needed to conduct TSSEM. Thus, it was decided that three would be the minimum number of studies needed for a TSSEM analysis in order for synthesised evidence to be meaningful. Mental health outcomes such as global psychopathological symptoms, depressive symptoms, anxiety symptoms, stress, and negative affectivity were chosen as the target psychological outcome to be used in TSSEM meta-analyses. Combining statistical synthesis with narrative summaries of mediation findings across studies will allow us to make stronger and more compelling conclusions regarding how MBIs improve psychological functioning and wellbeing, compared to using either of these techniques in isolation.

Table 1.

Theoretical models and proposed mechanisms of mindfulness and mindfulness-based interventions

Baer (2003)	Brown et al. (2007)	Grabovac et al. (2011)	Holzel et al. (2011)	Shapiro et al. (2006)	Vago & Silbersweig (2012)
Exposure; Cognitive change (decentering); Self-management (increased adaptive coping skills); Relaxation; Acceptance.	Insight (decentering); Exposure; Nonattachment (non-aversion and non-craving); Enhanced mind-body functioning; Integrated functioning (behaving in more purposeful ways).	Acceptance/ Compassion; Attention regulation; Ethical practices; Nonattachment and nonaversion; Mental proliferation (narrative thought processes).	Attention regulation; Body awareness; Emotion regulation (reappraisal, exposure, extinction, and reconsolidation); Change in perspective on the self.	Reperceiving (decentering); Self-regulation (stability of functioning and adaptability to change); Emotional, cognitive, and behavioural flexibility; Values clarification; Exposure.	Self-awareness; Self-regulation; Self-transcendence (self-other connection).

Note. Clarification of meanings are given in parentheses.

Method

Identification and Selection of Studies

A comprehensive search of published studies up to 10/01/2014 was conducted using the following electronic databases: PsycInfo, Scopus, Web of Knowledge, PsycArticles, ASSIA, and Science Direct. The search term was: *("mindfulness based cognitive therapy" OR "mindfulness based stress reduction" OR MBSR OR MBCT) AND (mechanism* OR mediat* OR predict* OR process* OR "structural equation modeling" OR caus* OR path* OR correlat* OR relationship OR associat*).*

Inclusion and Exclusion Criteria

Any study published in English which used 1) an adult sample (> 18 years), 2) a RCT or quasi-experimental design measuring pre to post-MBI change in variables, 3) a well-established method of mediation analysis with group (i.e. MBI versus control) as the independent variable, 4) MBSR or MBCT in the mediation analysis, 5) quantitative assessment (self- or other-report) of pre-post change in mental health and wellbeing outcomes (clinical and non-clinical) or related constructs, and 6) quantitative assessment of pre-post change in hypothesised psychological mediators, were included in the review. Studies which used adapted versions of MBSR and MBCT were included if the intervention followed the basic structure of MBSR and MBCT and involved both formal and informal mindfulness practice (meditation during sessions and for homework). The exclusion criteria were as follows: 1) uncontrolled studies, 2) reviews, 3) qualitative studies, 4) studies in which the MBI was not delivered in person (i.e., self-help format, internet-delivered MBIs), and 5) studies which tested mediators without a strong theoretical basis, or mediators not previously hypothesised in theoretical models and summaries of the mechanisms of mindfulness and MBIs. A strong theoretical basis for the study of variables as mediators is proposed by Kazdin (2007) to be a design requirement to establish mediation.

Quality Assessment

Mediation studies meeting the inclusion criteria were assessed for methodological quality and potential for bias using an appraisal framework adapted from Lubans et al. (2008), the CONSORT checklist (Schulz, Altman, & Moher, 2010),

and the Jadad checklist (Jadad et al., 1996), and informed by Kazdin (2007)'s design requirements for mediation. A score for each study was computed by assigning a value of 0 (no) or 1 (yes) to the 16 questions detailed in Tables 1 and 2 in Appendix B. If a study did not explicitly report information related to a certain question or if a question did not apply to a study, it was assigned 0 for that question. If a study with missing information explicitly reported that it was embedded in a larger trial, information provided in the original paper was also consulted to grade the study. Studies which scored 0-5 were classified as low-quality with a high risk of bias, 6-11 indicated a moderate risk of bias, and 12-16 high-quality with a low risk of bias.

Levels of Scientific Evidence

To draw narrative conclusions across included studies regarding the strength of evidence for identified mediators of MBIs, similar to previous reviews, a best evidence synthesis rating system (BESRS) was applied (Hoogendoorn, van Poppel, Bongers, Koes, & Bouter, 2000; Singh, Mulder, Twisk, & Chinapaw, 2008; van Stralen et al., 2011). Under this system, a body of evidence was considered strong if it involved consistent findings in two or more high-quality studies. Moderate evidence referred to consistent findings in one high-quality study and at least one low-quality study, or consistent findings in multiple low-quality studies, and insufficient evidence referred to when only one study was available or findings were inconsistent across two or more studies. Consistency was defined as at least 75% of studies demonstrating results in the same direction.

Data Extraction, Synthesis, and Statistical Analysis

The following data were extracted for narrative summaries of the mechanisms identified from included studies: study authors, year, sample characteristics, design, intervention, measurement times, measures, mediator(s), outcome(s), and type of mediation analysis. Additionally, in order to conduct TSSEM analyses for identified mechanisms tested with mental health outcomes, bivariate correlations between X (MBI versus control), change in M over the course of the intervention, and change in Y (clinical outcome) over the course of the intervention were extracted from each relevant study. Corresponding study sample sizes were also extracted. If a study did not explicitly report bivariate correlation coefficients, means, standard deviations, *t*-

statistics, *F*-statistics, and effect sizes were used to calculate bivariate correlations (Lipsey & Wilson, 2001). If a study provided insufficient data to calculate correlations, it was omitted from the TSSEM analyses.

A systematic approach was devised to select which mental health outcome to use from each relevant study, where studies measured more than one mental health outcome. The outcome used in the first instance was a global measure of psychopathology, then anxiety or depression. For studies which tested both anxiety and depression as outcomes, the outcome selected was the one which matched the study sample. For samples which did not match either anxiety or depression, the outcome used depended on whether baseline anxiety or depression levels in the sample were higher. If a study did not measure anxiety or depression, stress was chosen as the mental health outcome and failing that, negative affect. Finally, if two or more measures of the same outcome were used, the one with the stronger psychometric properties was selected. If a study did not include a mental health outcome, it was excluded from TSSEM. Only one mental health outcome was selected from each study to enable bivariate correlations to be extracted. Although it would be possible to calculate mean correlations across multiple outcomes in a single study, it would not be straightforward to determine the appropriate sampling variance of averaged correlations. Thus, it was deemed more appropriate to select one mental health outcome per study.

The metaSEM package (Cheung, 2013) in R (R Development Core Team, 2013) was used to perform the TSSEM analyses. The first stage tested the homogeneity of correlation matrices across studies and if these were not significantly different from each other, a pooled correlation matrix was produced. The second stage involved treating the pooled matrix as the observed correlation matrix and fitting a structural mediational model to the matrix to test the fit of model to the data. As samples, design, and effect sizes were expected to differ across studies, conducting random-effects models were deemed more appropriate than fixed-effects. Unstandardised regression coefficients and standard errors from the pooled matrix were then used to conduct Sobel tests, to determine the significance of the indirect pathway from participation in an MBI versus control to the mental health outcome, via particular mediators.

Results

Study Flow and Characteristics

The search identified 1,547 articles, 533 of which were duplicates. Of the remaining 1,014 papers, 845 were excluded based on the title or abstract, thus the number of full-text articles assessed for eligibility was 169. The final number of studies meeting the inclusion criteria and included in the systematic review was 20. One RCT (Batink et al., 2013) conducted two separate mediation analyses on subgroups in their sample and were included as two separate comparisons. Reference lists of full texts were examined for further studies meeting the inclusion criteria. Figure 1 illustrates the flow of studies through the review.

Tables 3 and 4 in Appendix B present summary data from the 20 included studies. Of these, 15 were RCTs, which compared MBCT to a non-active control (n = 6), active control (n = 2), or both (n = 1), or MBSR to a non-active control (n = 6), and 5 were quasi-experimental studies, which compared MBCT to a non-active control (n = 2), or MBSR to a non-active control (n = 2) or active control (n = 1). Study sample sizes ranged from 27 to 205 participants. Eight studies were conducted with adults with depressive symptoms, four used cancer patients or survivors, three used non-clinical samples, two used adults with heterogeneous anxiety disorders, two used adults with distress symptoms, and one used an unselected sample. Mental health outcomes (depression, anxiety, stress, distress, negative affectivity) were the most commonly assessed across studies (n = 18), with many studies including more than one mental health measure. Other outcomes assessed include mood states (n = 3), quality of life (n = 1), goal attainment (n = 1), and anger expression (n = 1). The most commonly tested mechanism across studies was mindfulness (n = 16), followed by rumination (n = 7), worry and concerns (n = 5), self-compassion (n = 3), psychological flexibility (n = 1), emotional reactivity (n = 1), cognitive reactivity (n = 1), and AMS (n = 1). Most studies performed more than one method of mediation analysis. Twelve studies applied the causal-steps test, 9 used bootstrapping, 6 reported the Sobel test, and 3 used Kraemer et al.'s (2002) framework for RCTs. Quality scores for included studies ranged from 5 (low-quality) to 14 (high-quality). Three studies were classified as low-quality, 15 were deemed to have a moderate risk of bias, and two studies were high-quality.

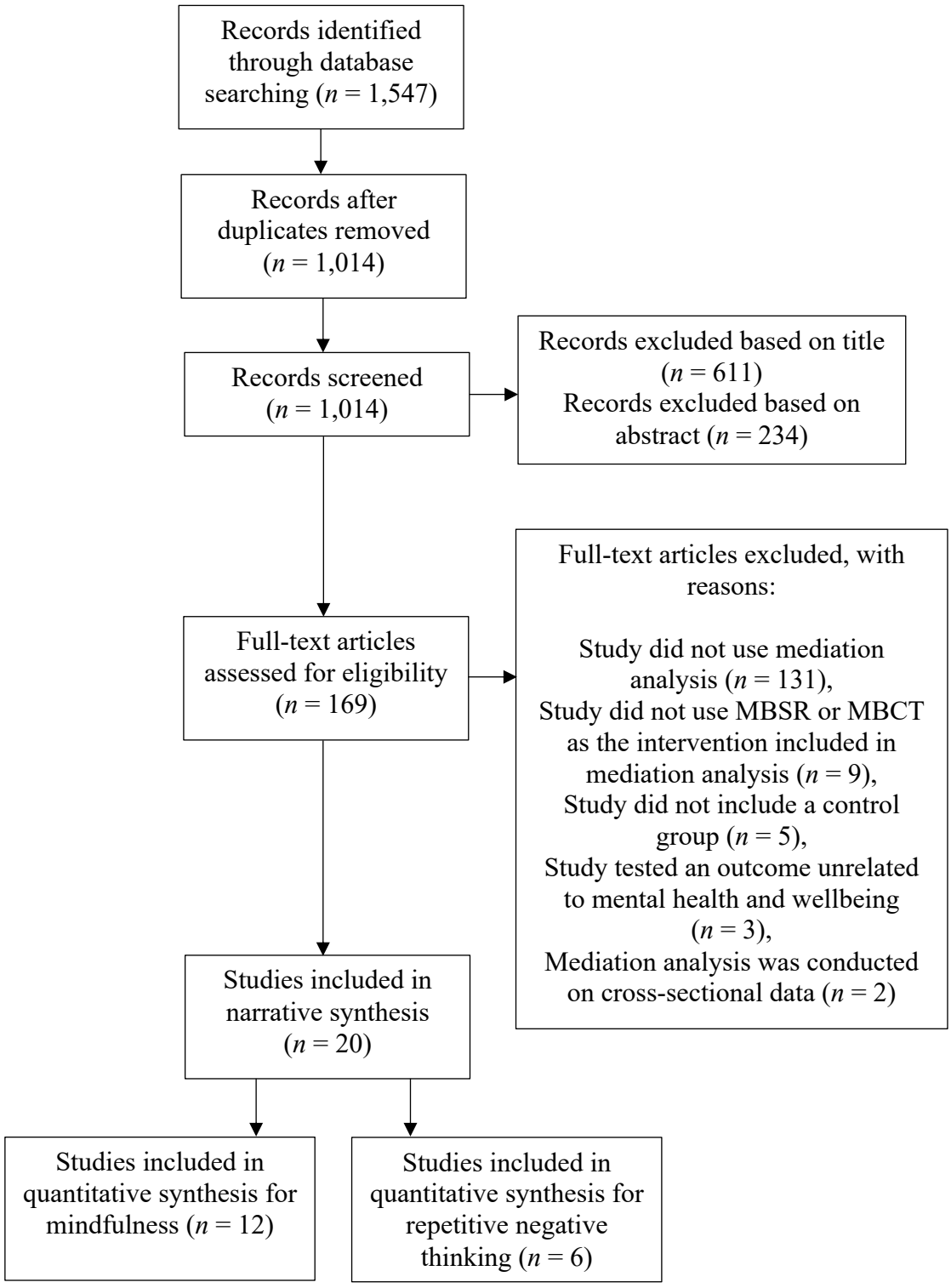


Figure 1. Flow of studies through the review process.

Summaries of Identified Mechanisms

This section gives narrative summaries of and evaluates the strength of evidence for each identified mechanism. Similar mechanisms (e.g., rumination and worry) are grouped together, as are mechanisms only tested in one study. This section also reports the results from TSSEM analyses for mediators tested by a substantial body of included studies (three or more) as mechanisms underlying the effects of MBIs on clinical outcomes.

Mechanism 1: Mindfulness. Perhaps the most important question about how MBIs work concerns whether or not their effects on psychological outcomes are mediated by increases in mindfulness, as theoretically asserted by proponents of MBCT and MBSR (Kabat-Zinn, 1982; Segal et al., 2002, 2013). In the 12 RCTs and 4 quasi-experimental studies which examined mindfulness as a mediator, most studies measured levels of depression (n = 9) as the main psychological outcome, followed by stress (n = 3), anxiety (n = 3), mood states (n = 3), quality of life (n = 1), and anger expression (n = 1), with many studies measuring more than one outcome. Samples included adults with depressive symptoms (n = 6), cancer (n = 3), distress symptoms (n = 2), and anxiety disorders or symptoms (n = 2). Non-clinical samples (n = 2) and an unselected sample (n = 1) were also used. Twelve studies examined mindfulness as a mechanism in simple mediator analyses and four analysed its effects in multiple mediator models, of which two reported both simple and multivariate mediation analyses. Multiple mediation involves simultaneously testing multiple variables as mediators in a single model.

Generally, findings showed support for mindfulness as a mediator, regardless of heterogeneity in measures of mindfulness, sample characteristics, and outcomes. As evidence for mindfulness as a mechanism was found in both simple and multiple mediator analyses, this suggests its mediating effects were present over and above the effects of other tested mediators, which included worry, rumination, and psychological flexibility. However, Labelle, Campbell, and Carlson's (2010) low-quality quasi-experimental study did not find mindfulness to mediate the impact of MBSR on depressive symptoms. This could be due to design limitations acknowledged by the authors, such as lack of randomisation of participants to conditions and not testing only participants who received an adequate dose of MBSR and are thus sufficiently familiar with the principles of the intervention. Additionally, Keng, Smoski, Robins, Ekblad, and Brantley (2012) did not find evidence for mindfulness as a mediator of MBSR's

effects on anger expression, in their multiple mediation analysis of both self-compassion and mindfulness using bootstrapping. This may have been due to mindfulness and self-compassion scores being significantly correlated and thus not empirically distinct constructs. Preacher and Hayes (2008) cautioned against testing overlapping constructs in multiple mediator models, as this compromises the significance of indirect effects.

Despite this, of the studies which found evidence for mindfulness as a mediator, one was high-quality, eleven medium-quality, and two low-quality, which constitutes moderate, consistent evidence for mindfulness as a mediator, according to the BESRS (e.g., Hoogendoorn et al., 2000). However, many studies have methodological limitations which future research should improve upon in order to provide strong evidence for mindfulness as a mediator. For example, only three studies compared an MBI to an active control group, one measured change in mediator before the outcome, none measured change in mediator during treatment, nine conducted mediation analysis only on participants who received an adequate dose of the intervention, and eleven used the most appropriate method of mediation analysis given study design and recommendations (e.g., MacKinnon et al., 2002).

TSSEM Results with Mindfulness as a Mediator. Of the 16 included studies which tested mindfulness as a mediator, 13 included a measure of a mental health outcome and sufficient information to calculate bivariate correlations. Two of these studies (Nyklicek et al., 2013; Nyklicek & Kujipers, 2008) used overlapping samples for their mediation analyses involving separate variables. Therefore, although both were included in the narrative synthesis, only the one with the higher quality score (Nyklicek et al., 2013) was included in the TSSEM. Thus, correlation matrices were extracted and synthesised from 12 studies testing mindfulness for the first stage of the TSSEM (Batink et al., 2013; Bränström et al., 2010; Kuyken et al., 2010; Labelle, 2012; Labelle et al., 2010; McManus et al., 2012; Nyklicek et al., 2013; Raes et al., 2009; Shahar, Britton, Sbarra, Figueredo and Bootzin, 2010; van Aalderen et al., 2012; Vøllestad et al., 2011). One study (Batink et al., 2013) conducted two separate mediation analyses on subgroups in their sample and were included as two comparisons. The most common mental health outcome across studies selected for TSSEM was depression (n = 7), followed by stress (n = 2), anxiety (n = 2), and negative affect (n = 1). The total pooled sample size was 1109. Table 2, which presents the pooled correlation coefficients for X, M, and Y, shows that all three correlations

were significant. The *Q* statistic for the homogeneity of effect sizes was also non-significant ($Q(24) = 27.09$, $p = .30$), indicating that the 12 correlation matrices were relatively similar and justifying their synthesis. Figure 2 displays the path diagram of the mediational model fitted to the data for stage two of the TSSEM analysis. Although the regression coefficient for path *c'* remained significant, it has reduced in size compared to the value of path *c*, which is indicative of partial mediation (Baron & Kenny, 1986). A Sobel test using correlation estimates for X and M, and M and Y, and their standard errors demonstrated that mindfulness significantly mediated the effects of MBIs on mental health outcomes ($z = 4.99$, $SE = 0.02$, $p < .001$).

Table 2.

Pooled correlation coefficients (k = 12) for X (participation in MBIs vs control), M (changes in mindfulness), and Y (changes in mental health outcome)

	X	M	Y
X	1		
M	0.34 (0.03)*	1	
Y	0.27 (0.03)*	0.36 (0.06)*	1

Note. *$p < .001$. Standard errors are displayed in parentheses.

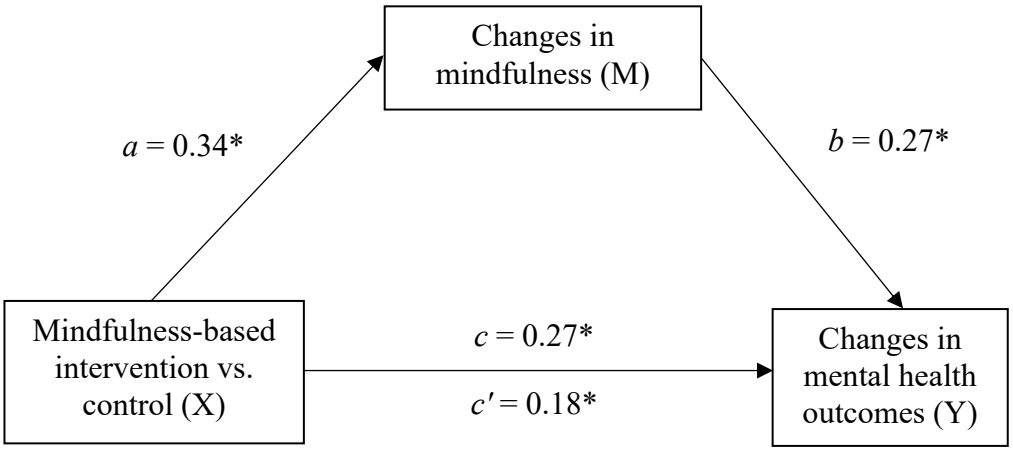


Figure 2. Path diagram depicting the stage two mediational model of the TSSEM, with changes in mindfulness as the mediator. Values are path coefficients. *Note.* *$p < .001$.

Mechanism 2: Repetitive Negative Thinking. RNT is a style of repetitive thinking about negative experiences which is difficult to disengage from and at least partly intrusive (Ehring et al., 2011). Two commonly examined forms of RNT include worry and rumination, which are typically highly correlated with each other (e.g., Fresco et al., 2002) and implicated in a range of psychopathologies (Ehring & Watkins, 2008). The only difference between worry and rumination is argued to be temporal orientation (Watkins, 2008), with worry relating more to the future and rumination more to the past (Papageorgiou & Wells, 1999). However, theory and evidence concerning the degree of overlap versus distinction between these constructs are inconsistent (e.g., Hoyer, Gloster, & Herzberg, 2009). For the purpose of this review, as findings generally suggest that rumination and worry appear to be closely related constructs (e.g., Watkins, Moulds, & Mackintosh, 2005), and there is a paucity of studies examining rumination and worry separately, both mechanisms are captured under RNT.

In the five RCTs and three quasi-experimental studies which have tested RNT constructs as mediators of MBIs, the most frequently assessed outcome was levels of depression ($n = 5$), followed by stress ($n = 2$), anxiety ($n = 1$), and global psychopathological symptoms ($n = 1$), with one study including measures of both stress and anxiety (Lengacher et al., 2014). Most studies recruited participants with depressive symptoms ($n = 4$), three used cancer patients or survivors, and one used a

non-clinical sample. Out of the eight studies, five examined RNT in simple mediation analyses and three in multiple mediator models, of which two reported both simple and multivariate mediator analyses. Other mediators included in multivariate analyses were mindfulness and psychological flexibility.

Despite differences in questionnaires, samples, and methods of mediation analysis conducted across studies, findings from all studies generally demonstrated that RNT constructs were significant and unique mediators of the effects of MBIs on clinical outcomes. Seven studies were deemed medium-quality, with a moderate risk of bias, and one study was low-quality with a high risk of bias, constituting consistent, moderate evidence for RNT as a mediator, according to the BESRS (e.g., Hoogendoorn et al., 2000). Only Shahar et al. (2010) found equivocal evidence for rumination as a mediator of MBCT's effects on depression in their multiple mediation analysis of both rumination and mindfulness using bootstrapping. They found that mindfulness and the brooding aspect of rumination, which refers to the tendency to dwell on negative thoughts related to one's condition, were significant mediators, but no evidence for the reflective pondering component, which is concerned with constructively thinking about one's condition in order to better understand and improve it (Treynor, Gonzalez, & Nolen-Hoeksema, 2003). This may be due to brooding being the key maladaptive component in rumination rather than reflection, consistent with findings from research which showed that brooding but not reflection mediated the relationship between negative thinking and depression (Lo, Ho, & Hollon, 2008) and self-criticism and suicidal ideation (O'Connor & Noyce, 2008).

Although there appears to be a consistent body of evidence for RNT as a unique mechanism underlying the effects of MBIs, included studies contain numerous methodological limitations which preclude strong conclusions regarding mediation and specificity of improvements to MBIs. For example, none of the eight studies compared an MBI with an active control group and measured change in mediator before the outcome or during treatment, only three conducted mediation analysis on participants who received an adequate dose of the intervention, and only four used the most appropriate method of mediation analysis. Future research examining RNT constructs as mediators would be enhanced by taking such limitations into account.

TSSEM Results with RNT as a Mediator. Six of the eight included studies which tested RNT constructs as a mediator included a measure of a mental health outcome and sufficient information to calculate bivariate correlations and were

included in the second TSSEM analysis (Heeren & Philippot, 2011; Labelle, 2012; Labelle et al., 2010; Lengacher et al., 2014; Shahar et al., 2010; van Aalderen et al., 2012). The most common mental health outcome across studies selected for TSSEM was depression (n = 3), followed by stress (n = 1), anxiety (n = 1), and global psychopathological symptoms (n = 1). The total pooled sample size was 586. All three pooled coefficients were significant (see Table 3) and the Q statistic was non-significant, indicating homogeneity of correlation matrices, $Q(11) = 18.79$, $p = .07$. As depicted in Figure 3, after accounting for M, the unstandardised coefficient for path *c'* reduced in size compared to the value for path *c* but remained significant, indicating partial mediation (Baron & Kenny, 1986). A Sobel test showed that RNT significantly mediated the effects of MBIs on mental health outcomes ($z = 4.88$, $SE = 0.02$, $p < .001$).

Table 3.

Pooled correlation coefficients (k = 6) for X (participation in MBIs vs control), M (changes in repetitive negative thinking), and Y (changes in mental health outcome)

	X	M	Y
X	1		
M	0.31 (0.04)*	1	
Y	0.31 (0.06)*	0.33 (0.05)*	1

Note. *$p < .001$. Standard errors are displayed in parentheses.

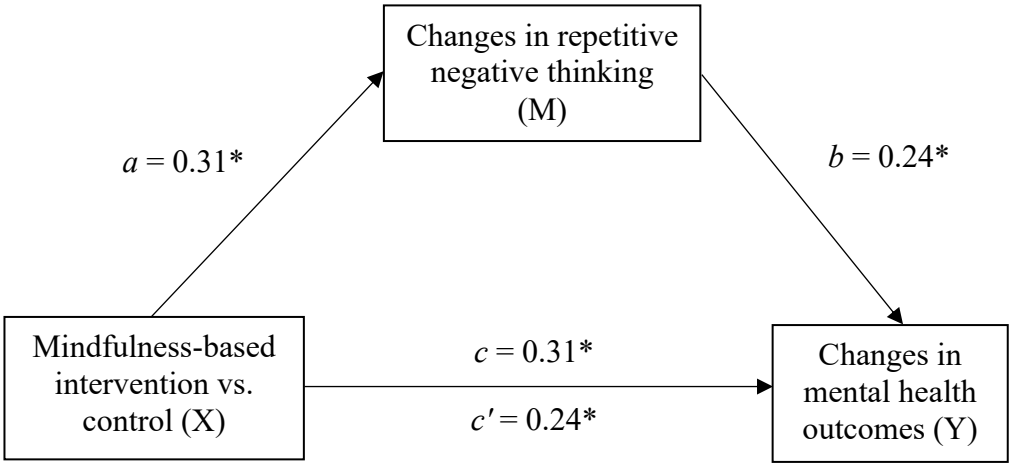


Figure 3. Path diagram depicting the stage two mediational model of the TSSEM, with improvements in repetitive negative thinking (worry and rumination) as the mediator. Values are path coefficients. *Note.* *$p < .001$.

Mechanism 3: Self-compassion. Two RCTs and one quasi-experimental study tested self-compassion as a mediator of MBIs. One study assessed self-compassion as a mediator of an MBI's effect on depression in a sample of adults with recurrent depression (Kuyken et al., 2010) and the other two, on trait anxiety (Bergen-Cico & Cheon, 2013) and anger expression (Keng et al., 2012) in non-clinical samples. Keng and colleagues' study examined self-compassion in a multivariate mediation analysis, controlling for mindfulness; the other two studies reported simple mediation analyses.

Despite differences in samples and types of mediation analysis conducted, findings from Kuyken et al.'s (2010) study, with a low risk of bias, supported self-compassion as a mediator of MBCT's effects. However, Keng et al.'s (2012) study, with a moderate risk of bias, and Bergen-Cico and Cheon's (2013) quasi-experimental study, which had a moderate risk of bias, found that although MBSR participation significantly increased self-compassion, this increase did not mediate MBSR's effects on anger expression or anxiety. This could be due to reasons such as Keng et al.'s study only examining self-compassion in a multiple mediator analysis, the cultivation of self-compassion to improve anger expression and anxiety requiring a longer period of mindfulness practice compared to other outcomes, methodological limitations of the studies, and the presence of moderated mediation effects (Preacher, Rucker, & Hayes, 2007), or unmeasured moderators affecting the strength of the mediated relationship.

As evidence for self-compassion was found in just one high-quality study, this constitutes preliminary but inconsistent evidence for self-compassion as a mediator of the impact of MBIs on psychological outcomes (e.g., Hoogendoorn et al., 2000). This modest body of evidence could be extended and improved upon by including more studies which measure change in mediator before the outcome and during treatment, include an active control group, conduct mediation analysis only on participants who receive an adequate dose of the intervention, and use the most appropriate method of mediation analysis. Methodological limitations, coupled with the small number of studies which supported self-compassion as a mediator, means that we are unable to establish self-compassion as a mechanism of MBIs. Furthermore, as only two of the three mediation studies which tested self-compassion as a mediator included a measure of a mental health outcome, combining findings using TSSEM was deemed inappropriate at this time.

Mechanism 4: Cognitive and Emotional Reactivity. Two high-quality RCTs, with a low risk of bias, studied cognitive and emotional reactivity as mediators of MBCT's effects on depressive symptoms, using samples of adults with recurrent depression. Cognitive and emotional reactivity refer to the extent to which a mild state of distress coupled with stress reactivates negative thinking and emotional patterns, putting individuals at risk of a depressive episode (Scher, Ingram, & Segal, 2005). One study found evidence for improvement in emotional reactivity as a mediator of MBCT's effects on depressive symptoms in a simple mediation analysis (Britton, Shahar, Szepsenwol, & Jacobs, 2012). However, Kuyken et al. (2010) found that post-treatment cognitive reactivity was in fact greater in the MBCT group compared to the control group, and that greater reactivity only predicted increased depression (measured 12 months after the MBCT programme) for control group participants.

Other Mechanisms: Psychological Flexibility and Autobiographical Memory Specificity. Psychological flexibility refers to the ability to fully embrace thoughts, feelings, and experiences in the present moment without avoidance, and persisting or altering behaviour to be consistent with goals and values (Hayes, Luoma, Bond, Masuda, & Lillis, 2006). A medium-quality quasi-experimental study which used a sample of cancer patients (Labelle, 2012) found psychological flexibility to be a significant mediator of MBSR's effects on stress and mood states in both simple and multiple mediator models, with RNT taken into account. According to the BESRS (e.g., Hoogendoorn et al., 2000), this constitutes preliminary but insufficient evidence

for psychological flexibility as a mechanism. AMS was not found to be a significant mediator of MBCT's effects on likelihood of goal attainment in a low-quality RCT (Crane et al., 2012) using a sample of adults with residual depressive symptoms. This non-significant finding may have been due to the high risk of bias associated with Crane et al.'s study and methodological limitations, such as not measuring change in mediators during the intervention and before outcomes, not conducting mediation analysis on only participants who received an adequate dose of MBCT, and not using the most appropriate method of mediation analysis.

Discussion

This paper's aims were threefold. First, it aimed to systematically review mediation studies in the literature on MBIs, in order to identify the mechanisms underlying MBCT and MBSR's effects on psychological functioning and wellbeing. Second, it aimed to evaluate and provide narrative summaries of the strength of evidence for each identified mechanism. Third, it aimed to conduct TSSEM analyses, to statistically synthesise evidence for mechanisms of the effects of MBIs on mental health outcomes which were supported by three or more mediation studies. In the 20 included studies, the following psychological constructs were identified and tested as mediators of MBIs: mindfulness, repetitive negative thinking (rumination, worry), self-compassion, psychological flexibility, emotional reactivity, cognitive reactivity, and AMS. The narrative synthesis described consistent and moderate evidence for mindfulness and RNT, and preliminary but insufficient evidence for cognitive and emotional reactivity, self-compassion, and psychological flexibility as mechanisms of the effects of MBIs on clinical and non-clinical psychological outcomes. The Sobel tests in two separate TSSEM analyses demonstrated that both mindfulness and RNT significantly mediated the effect of MBIs on mental health outcomes such as anxiety, depression, global psychopathological symptoms, stress, and negative affect. Although in both TSSEM analyses the causal-steps test indicated partial mediation, several researchers have criticised the distinction between full and partial mediation as being trivial and emphasised the importance of avoiding these concepts when interpreting mediation (e.g., Rucker, Preacher, Tormala, & Petty, 2011). Additionally, unlike the causal-steps method, the Sobel test is able to assess whether or not the indirect,

mediated effect is statistically significant (MacKinnon et al., 2002). Therefore, current TSSEM findings are better interpreted using Sobel test over causal-steps results.

These findings are largely consistent with the theoretical underpinnings of MBSR and MBCT. Evidence for mindfulness as a mechanism supports the key theoretical premise underlying MBSR and MBCT that the cultivation of mindfulness skills leads to insight into and acceptance of one's experience (Kabat-Zinn, 1982; Segal et al., 2002, 2013), which in turn leads to positive outcomes. Evidence for rumination and worry as mechanisms is also reflected in the underlying theory of MBCT, participation in which is postulated to decrease depressive recurrence through increasing insight into and disengagement from recurrent maladaptive thinking about one's depressive symptoms (Segal et al., 2002, 2013).

Current findings also offer suggestions for theory development. For instance, models explaining how mindfulness and MBIs work could incorporate self-compassion more explicitly as a mechanism. Theoretical accounts of mindfulness and MBIs in combination with this systematic review also highlight gaps in the literature and inform future research assessing the mediators of MBIs. For example, although theories propose attentional control and regulation (e.g., Holzel et al., 2011; Shapiro et al., 2006), bodily awareness (Holzel et al., 2011), self-awareness, self-regulation, and self-transcendence (Vago & Silbersweig, 2012) as key mechanisms which work alongside empirically-tested mechanisms, this review did not identify any studies which have tested these variables as mediators of MBIs. As well as informing theory and future research, current findings also have implications for clinical practice. For instance, they suggest that we may be able to develop MBIs to maximise their therapeutic effects, through emphasising components of these interventions related to mindfulness, rumination, worry, cognitive and emotional reactivity, self-compassion, and psychological flexibility.

Although integrating the findings across the mediation studies in this review advances our understanding of the causal pathways between MBIs and psychological outcomes and guides future research, theoretical developments, and clinical practice, most included studies had a moderate risk of bias and at least one of several key methodological shortcomings. For example, although temporal ordering of mediator and outcome variables is crucial to establishing mediation (Kazdin, 2007), most studies did not conduct mediation analyses that took temporality into account, thus weakening conclusions about causality. Most studies also lacked active control conditions and did

not conduct mediation analysis only on participants who received an adequate dose of the MBI, meaning that change cannot be readily attributed to the MBI rather than to non-specific factors (e.g., participant expectation of benefit, attention from research team, group support). Therefore, despite their worth, in the presence of these design flaws, the current body of evidence is likely to be subject to biases and insufficient to provide compelling support for the identified mechanisms. Once future research improving on these limitations has established the mechanisms of MBIs, a next step could be to examine each mechanism in depth, to explore the temporal ordering of its effects in relation to other mechanisms in a multiple mediator model, the degree of overlap between mechanisms in multiple mediator models, the impact of moderator variables on the mediated relationship, and whether experimental manipulation of the mechanism affects levels of the outcome.

Strengths and Limitations of the Current Review

As no known studies have systematically reviewed and provided narrative and statistical syntheses of findings from mediation studies exploring the mechanisms underlying MBIs, the main strength of the current study lies in addressing this gap in the literature. Although Chiesa, Anselmi, and Serretti (2014) recently published a narrative review of the literature on psychological mechanisms of MBIs, their review did not document a systematic, replicable literature search strategy, examined case control and uncontrolled studies alongside RCTs, the former of which do not allow us to infer causality, did not specifically examine mediation studies, a crucial analytic method for establishing a variable as a mediator, and did not conduct statistical syntheses of findings across studies. Another strength of this systematic review is that for the purpose of the narrative synthesis, it did not confine its search to studies with specific samples and outcomes, reflecting the perspective supported by many proponents of mindfulness, that the underlying mechanisms of its effects are the same across different intensities of suffering in all people (e.g., Teasdale & Chaskalson, 2011). The non-significant homogeneity statistics associated with TSSEM analyses, which indicate that the roles of mindfulness and rumination and worry as mechanisms did not significantly differ across different samples, appear to support this approach. However, it is possible that some mechanisms are unique to particular populations or play a larger role in explaining therapeutic change for certain populations than others.

Due to the dearth of mediation studies of MBIs, the current review did not examine population-specific mechanisms as this would have further restricted the inclusion criteria. As the number and quality of mediation studies increase, it would be worthwhile for future reviews to address population-specific mechanisms of MBIs.

A further strength of this review is that it only included studies which used MBCT or MBSR as the MBI, as other interventions which include teachings in mindfulness such as ACT (Hayes & Wilson, 1994) and DBT (Linehan, 1993) differ from MBCT and MBSR and each other in the duration and frequency of mindfulness practice involved, the inclusion of mindfulness psychoeducation and non-mindfulness therapeutic components, and a consensus has not yet been reached regarding the similarity between MBCT and MBSR and these alternative interventions (Chiesa & Malinowski, 2011). Inclusion of multicomponent interventions could inflate effect sizes and make it unclear whether mindfulness or other components lead to positive changes. Thus, including only MBSR and MBCT minimises methodological heterogeneity across studies, allowing meaningful narrative summaries to be constructed and justifying integration of studies using TSSEM. However, it is worth noting that some of the identified mechanisms in this review may also account for the therapeutic effects of alternative interventions. In particular, psychological flexibility has been explored in relation to ACT and found to mediate the effects of ACT on a range of outcomes such as emotional exhaustion (Lloyd, Bond, & Flaxman, 2013), pain-related disability, and life satisfaction (Wicksell, Olsson, & Hayes, 2010). Thus, by excluding mediation studies of alternative interventions which include mindfulness teachings, as well as studies of other closely related therapies (e.g., mindfulness-based relapse prevention; Bowen, Chawla, & Marlatt, 2010), this review may have underemphasised the role processes such as psychological flexibility play in implementing therapeutic change. A separate systematic review of mediation studies of alternative interventions would shed light on whether underlying processes are similar across different interventions with mindfulness components.

A limitation of this systematic review is that its conclusions may be affected by between-study differences in design. As the number of publications on mediation studies increases, it will be useful to examine only RCTs, in order to base conclusions on more rigorously designed studies and decrease heterogeneity in methods and quality across studies. As the number of mediation studies increases, other potential limitations, such as rumination and worry being subsumed under RNT rather than

assessed separately can also be addressed. Included studies also used a range of measures of mechanisms and outcomes, exposing the review to measurement heterogeneity. However, all included studies were published in peer-reviewed journals and all measures had good psychometric properties. An additional limitation of this review pertains to its strict inclusion criteria, which excluded studies examining mechanisms without a strong theoretical basis, in line with Kazdin's (2007) design requirement to establish mediation. In doing so, it may have overlooked some mechanisms of MBSR and MBCT which have the potential to inform theoretical developments, such as positive and negative affect (Batink et al., 2013). Future reviews could improve on this limitation by instead including this design requirement as a quality criterion.

Furthermore, it is possible that the findings from this review are subject to the influence of publication bias. This applies to meta-analyses more generally (Turner, 2013). Although a few included studies reported non-significant indirect effects from their mediation analyses, it is possible that current findings overestimate the effects of the identified mechanisms in mediating the relationship between MBIs and outcomes. For example, some RCTs of MBCT or MBSR may have included rumination as a measure but not published mediation analyses. This may be due to mediation analyses being conducted post-hoc and thus only being reported when there is supporting evidence. If this is the case, there may be a greater risk of publication bias for mediation studies from RCTs than for RCT outcome studies, and this should be taken into account when interpreting the findings of this review.

Nevertheless, our review presents a valid synthesis of publicly available findings in the field and the aggregated data (AD) meta-analytic approach taken for the TSSEM analyses is typically used to inform practice and policy. Although an individual participant data (IPD) approach, in which raw data for each participant from each study are used for synthesis, has been advocated as a less biased method (Stewart & Parmar, 1993), currently the best strategy would be to conduct AD synthesis of findings before IPD meta-analysis due to the greater cost of IPD analysis and lack of available IPD datasets (Cooper & Patall, 2009). As the current review was the first to narratively and statistically synthesise findings from mediation studies exploring the mechanisms underlying MBCT and MBSR, the AD approach was taken using published results. However, future systematic reviews of this kind should consider the merits of implementing the IPD meta-analytic approach. Additionally, good practice

for future RCTs aiming to conduct mediation analyses would be to publish their full trial protocol and analytic plan prior to analysing their data. It is worth noting that current conclusions are not informed merely by the number of significant findings across studies and significant TSSEM results. They are based to a larger extent on whether or not individual studies are adequately designed to establish mediation, the importance of which is reiterated throughout this paper. Therefore, the cautious approach to evaluation taken in this review should minimise the impact publication bias has on the validity of its conclusions.

Conclusions

The empirical investigation of the mechanisms of change underlying the effects of MBIs on psychological functioning and wellbeing is a complex yet crucial path on which to embark in order to improve the quality, delivery, and effectiveness of the interventions, develop the theoretical underpinnings of mindfulness and MBIs, and inform the direction of future research. The current study was the first to systematically review mediation studies to identify and evaluate the strength and consistency of evidence for mechanisms underlying the effects of MBIs. It was also the first to use TSSEM analyses and accompanying Sobel tests to statistically synthesise evidence across mediation studies testing the effects of MBIs on mental health outcomes. It found moderate, consistent evidence for mindfulness and RNT, and preliminary but insufficient evidence for cognitive and emotional reactivity, self-compassion, and psychological flexibility as mechanisms. Moreover, TSSEM analyses and Sobel tests demonstrated evidence for mindfulness and RNT as significant mediators of the impact of MBIs on clinical outcomes. Although included studies using mediation analysis have key methodological shortcomings which preclude strong conclusions regarding mediation, they provide valuable insights into the potential causal pathways connecting MBIs with improved psychological outcomes and construct important groundwork on which future theories and research could build.

Chapter 3:
Investigating the Specific Effects of an Online Mindfulness-Based Self-Help Intervention on Stress and Underlying Mechanisms

Chapter 3 is published in *Mindfulness* as:

Gu, J., Cavanagh, K., & Strauss, C. (2017). Investigating the Specific Effects of an Online Mindfulness-Based Self-Help Intervention on Stress and Underlying Mechanisms. *Mindfulness*, 1-13.

Author Contributions

Design:	All authors
Data Collection:	JG
Data Analysis:	JG
Initial Draft:	JG
Reviewing and Editing:	All authors

Abstract

Previous research examining the effects of mindfulness-based interventions (MBIs) and their mechanisms of change has been hampered by failure to control for non-specific factors, such as social support and interaction with group members, facilitator contact, and expectation of benefit, meaning that it remained possible that benefits of MBIs could have been attributable, perhaps entirely, to non-specific elements. This experimental study examined the effects of a two-week online mindfulness-based self-help (MBSH) intervention compared to a well-matched classical music control condition and a waitlist control condition on perceived stress. This study also tested mindfulness, self-compassion, and worry as mechanisms of the effects of MBSH versus both control conditions on stress. University students and staff ($N = 214$) were randomised to MBSH, classical music, or waitlist conditions and completed self-report measures pre-, mid-, and post-intervention. Post-intervention, MBSH was found to significantly reduce stress compared to both control conditions. Bootstrapping-based mediation analyses used standardised residualised change scores for all variables, with mediators computed as change from baseline to mid-intervention, and the outcome computed as change from baseline to post-intervention. Changes in mindfulness, self-compassion, and worry were found to significantly mediate the effects of MBSH versus both control conditions on changes in stress. Findings suggest that cultivating mindfulness specifically confers benefits to stress and that these benefits may occur through improving theorised mechanisms.

Keywords: mindfulness, self-compassion, stress, mechanisms, self-help, online, RCT

Introduction

Mindfulness is commonly defined as the quality of awareness that arises through intentionally observing the stream of moment-to-moment experience in an open, accepting, and non-judgemental way (Kabat-Zinn, 1994). Over the past few decades, cultivating mindfulness through mindfulness practice has received increasing attention. Much of this can be attributed to the development of mindfulness-based interventions (MBIs) in clinical contexts, of which the most widely employed are mindfulness-based stress reduction (MBSR; Kabat-Zinn, 1982) and mindfulness-based cognitive therapy (MBCT; Segal, Williams, & Teasdale, 2002, 2013). MBSR and MBCT are eight-session group-based interventions that teach participants mindfulness through a range of mindfulness practices and teacher-led discussion, with the intention of improving wellbeing and mental health. This intention is supported by findings from randomised controlled trials (RCTs). Compared to control conditions, MBIs are effective at improving a range of outcomes, such as quality of life, severity of anxiety and depressive symptoms, risk of depressive relapse, stress, and chronic pain (e.g., Chiesa & Serretti 2009; Godfrin & Heeringen, 2010; Green & Bieling, 2012; Grossman, Tiefenthaler-Gilmer, Raysz, & Kesper, 2007; Hofmann, Sawyer, Witt, & Oh, 2010; Kuyken et al., 2016; Strauss, Cavanagh, Oliver, & Pettman, 2014).

In addition to the evidence base for the effectiveness of MBIs on clinically-relevant outcomes, studies have started to investigate the mechanisms underlying their effects using mediation analysis, which examines the indirect effect of a treatment (X) on an outcome (Y) through a mediator (M), or intervening variable. Kazdin (2007) described a number of benefits of identifying how psychotherapies work, including the potential to better understand the outcomes of treatments, enhance aspects of interventions to optimise therapeutic benefits, facilitate the translation of research on treatments into practice, and identify treatment moderators so that therapies can be matched to individuals. Recently, Gu, Strauss, Bond, and Cavanagh (2016) systematically reviewed mediation studies examining the effects of MBSR and MBCT compared to control conditions on mental health and wellbeing outcomes and evaluated the strength of evidence for each identified mechanism. They found that most mediation studies selected mediators based on the theoretical underpinnings of these MBIs, with the most commonly tested mechanism being mindfulness, followed by repetitive negative thinking constructs (worry and rumination) and self-compassion.

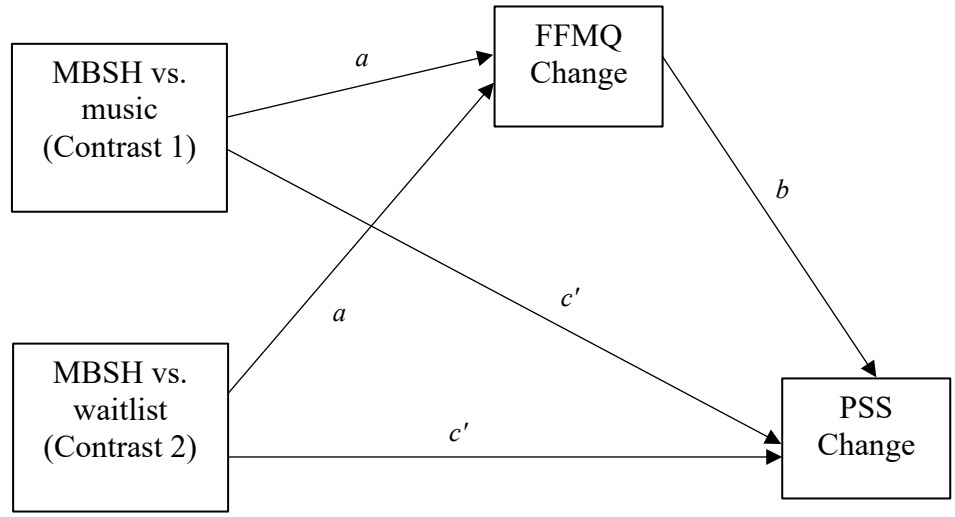
They identified moderate and consistent evidence for mindfulness and repetitive negative thinking as mechanisms, but insufficient evidence for self-compassion. However, many of the studies reviewed had at least one key methodological limitation. For example, most studies did not compare an MBI to a well-matched control condition. Comparing interventions to matched control conditions, which hold constant all factors except for the core, specific elements of the intervention, is important because this makes it possible to determine whether the specific elements of the intervention are responsible for beneficial outcomes (Mohr et al., 2009). Without adequately matched control conditions, it is possible that benefits are the result of non-specific elements of the intervention (e.g., expectation of benefit, facilitator contact, social support or interaction with group members) rather than specific elements (i.e., learning mindfulness).

Given the evidence base for the effectiveness of MBIs and emerging evidence illuminating their mechanisms, researchers have started to examine ways to increase the accessibility of MBIs to benefit more people. One way of extending the accessibility of interventions is to develop self-help versions (i.e., learning mindfulness through online courses, self-help books, smartphone applications, etc.). An additional benefit of mindfulness-based self-help (MBSH) interventions is that they remove many of the non-specific factors found in group-based MBIs, such as facilitator contact and group support and interaction, making them particularly suited to examining the effects of learning mindfulness specifically. Meta-analyses of RCTs of mindfulness-based self-help (MBSH) interventions and acceptance-based self-help interventions have shown beneficial effects on stress with a medium effect size, and depression, anxiety, wellbeing, and mindfulness with small to medium effect sizes (Cavanagh, Strauss, Forder, & Jones, 2014; Spijkerman, Pots, & Bohlmeijer, 2016). However, these findings were based largely on combining data from studies which compared MBSH interventions to waitlist control conditions. If we truly want to examine the effects of learning mindfulness and the specific change mechanisms associated with these effects, we not only want an MBSH intervention that removes non-specific elements of face-to-face MBIs, we would also want a matched control intervention that controls for additional non-specific factors (Mohr et al., 2009), namely, expectation of benefit and engagement with a structured intervention of a similar format and with similar time demands. To our knowledge, only nine studies have compared MBSH interventions to matched control conditions (Carissoli, Villani, & Riva, 2015; Dowd et al., 2015;

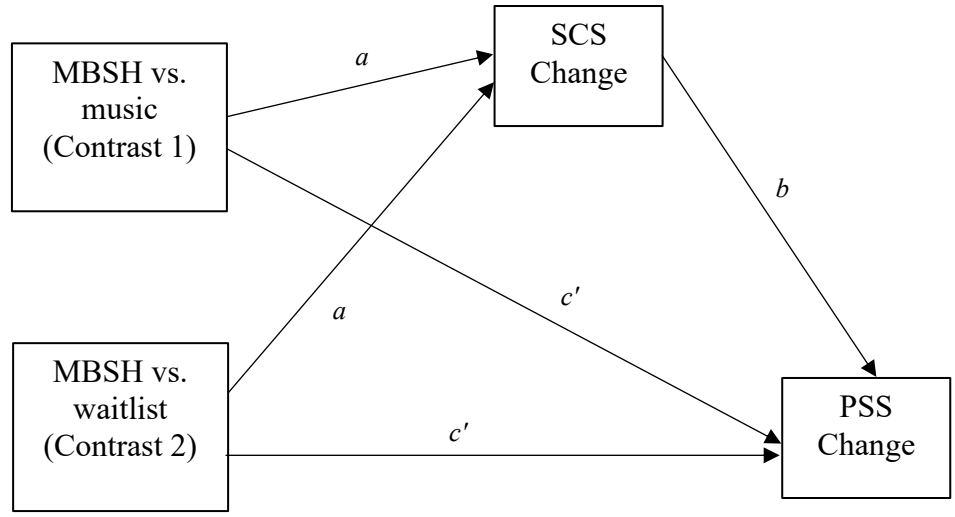
Howells, Ivtzan, & Eiroa-Orosa, 2014; Jimenez, 2008; Ly et al., 2014; Mongrain, Komeylian, & Barnhart, 2016; Niles et al., 2012; Stankovic, 2015; Wahbeh, Goodrich, & Oken, 2016). Of these studies, four examined online MBSH interventions, three audio CDs, and two smartphone applications, and four compared MBSH to psychoeducation, two to relaxation exercises, one to behavioural activation, one to list making, and one to expressive writing. None of these studies examined the mechanisms underlying the effects of learning mindfulness.

The primary aim of the present study was to examine whether a two-week online MBSH intervention significantly reduces perceived stress compared to an inactive waitlist control condition and a matched non-mindfulness condition. The matched control condition consisted of a two-week online classical music listening intervention, which shares the same structure, format, and time demands as the MBSH intervention. We also aimed to test three theoretically and/or empirically supported mechanisms of MBIs (Gu, Strauss, Bond, & Cavanagh, 2016) – namely, mindfulness, self-compassion, and worry – as mediators of the effects of MBSH versus the waitlist control and matched non-mindfulness conditions on stress. This study therefore investigates not only whether the specific process of learning mindfulness reduces stress, by comparing MBSH to a matched non-mindfulness condition and controlling for non-specific factors, but also how learning mindfulness specifically may reduce stress, by examining the three most commonly tested, and theorised, mechanisms of MBIs as mediators. This study tests three hypotheses. First, compared to the waitlist control condition, the MBSH intervention was predicted to significantly reduce stress over the two-week time period. No hypothesis was made about the relative effect of the MBSH versus classical music intervention on stress because there are no compelling theoretical or empirical reasons to hypothesise a difference between these conditions in either direction. Second, the effects of the MBSH intervention versus waitlist control on changes in stress were expected to be significantly mediated by changes in mindfulness, self-compassion, and worry (Figures 4a, 4b, and 4c). Third, the effects of the MBSH intervention versus the classical music intervention on changes in stress were hypothesised to be significantly mediated by changes in mindfulness, self-compassion, and worry (Figures 4a, 4b, and 4c).

a


MBSH vs.
music
(Contrast 1)
MBSH vs.
waitlist
(Contrast 2)
FFMQ
Change
PSS
Change
a
a
b
c′
c′

b


MBSH vs.
music
(Contrast 1)
MBSH vs.
waitlist
(Contrast 2)
SCS
Change
PSS
Change
a
a
b
c′
c′

c

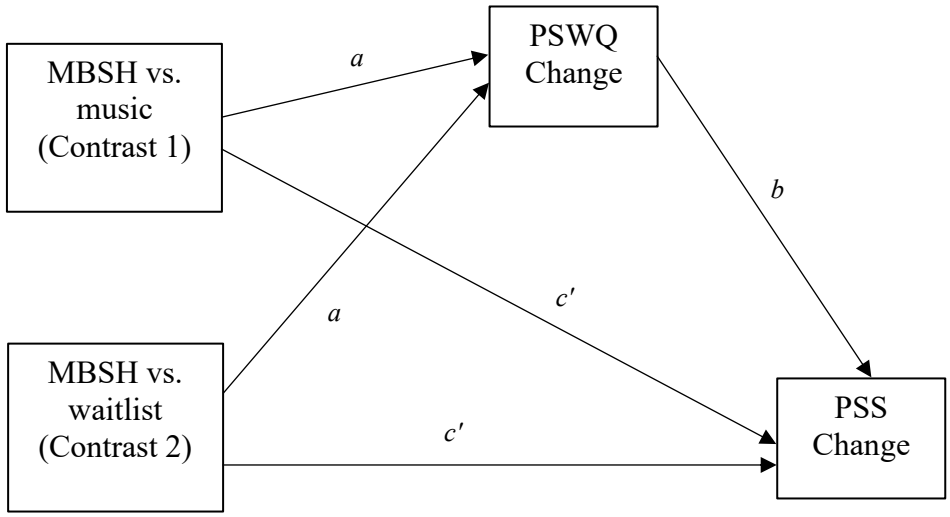


Figure 4. Path diagrams depicting three multicategorical independent variable mediation models. **a** The effects of mindfulness-based self-help (MBSH) versus music or MBSH versus waitlist control (X) on changes in perceived stress (PSS) (Y) mediated by changes in mindfulness (FFMQ) (M). **b** The effects of MBSH versus music or MBSH versus waitlist control on changes in perceived stress mediated by changes in self-compassion (SCS). **c** The effects of MBSH versus music or MBSH versus waitlist control on changes in perceived stress mediated by changes in worry (PSWQ). *a, b,* and *c'* are unstandardised regression coefficients which represent predicting M from X (*a*), Y from M controlling for X (*b*), and Y from X controlling for M (*c'*). The product of the *a* and *b* paths, *ab*, represents the mediated, or indirect effect.

Method

Participants

This study had an experimental design, testing the effects of group (MBSH, classical music, and waitlist control) and time (baseline, mid-intervention, and post-intervention) on self-reported levels of mindfulness, worry, self-compassion, and perceived stress. Participants were 214 students and staff (72.90% female) at a university in the South of England, with access to the university's online learning portal. Their ages ranged from 18 to 49 years (M = 24.20, SD = 5.79). Participant baseline demographic characteristics are presented in Table 4.

Procedure

Participants completed baseline questionnaires hosted on Bristol Online Surveys (www.onlinesurveys.ac.uk) and were randomised to the MBSH, classical music, or waitlist control condition. Randomisation was conducted by researchers independent of the research team and blind to participant details using a computer-generated blocked allocation method, with six numbers per block. Participants allocated to MBSH and music conditions were given access to the intervention sites within 24 hours of randomisation. Standardised e-mails were sent during the two-week period (days 3, 7, and 10) to encourage participants to engage in the interventions. Participants were also e-mailed at days 7 (mid-intervention) and 14 (post-intervention) to complete online questionnaires. Reminder e-mails for participants to complete mid-intervention and post-intervention measures were sent once for mid-intervention measures and three times for post-intervention measures.

Participants randomised to the waitlist control condition were sent standardised e-mails within 24 hours informing them of their allocation, and at days 7 and 14, asking them to complete mid- and post-intervention questionnaires. Due to limited study resources, it was not possible for researchers to be blind to condition allocation, but e-mail text was standardised to ensure that completion of questionnaires was not influenced by researcher bias. Participants were debriefed and given access to both online interventions upon completion of post-intervention questionnaires or after the two-week waiting list period.

The Online Mindfulness and Music Interventions. The MBSH intervention, 'Learning Mindfulness Online', was taken from Cavanagh et al. (2013). The classical music intervention, 'Listening to Classical Music Online', mirrored the structure and format of the MBSH intervention as closely as possible, but differed in content. Both interventions were hosted on the university's online learning portal.

Learning mindfulness online. This site consisted of eight pages. The welcome page provided information on what to expect from the intervention. *What is Mindfulness?* gave an overview of mindfulness and its benefits, history, and practice using text and a brief video clip. *Daily Mindfulness Practice* contained a ten-minute audio recording of a guided mindfulness practice and invited participants to listen to this for the duration of the intervention. Two versions of the recording were uploaded; one delivered by a female voice and one by a male voice, and participants could select which one they preferred to listen to. The audio clips were recorded by clinical psychologists trained to deliver MBIs. *Everyday Mindfulness Activities* described daily informal mindfulness practices. During the first week, participants were invited to bring mindfulness to one routine activity (e.g., showering) and during the second week, participants were invited to be mindful during a 5 to 30-minute walk. The *Daily Practice and Everyday Mindfulness Activities FAQ* page provided advice on how to approach commonly encountered experiences during mindfulness practice (e.g., boredom, discomfort, sleepiness). Participants could use the *My Daily Journal* page to record their thoughts and feelings related to mindfulness practice. The *Study Information* page contained study information and *Help and Assistance* gave the contact details of the research team, the university's counselling service, and local and national mental health services.

Listening to classical music online. This control site included eight matched pages; the welcome page, *Why Listen to Classical Music?, Daily Classical Music Listening, Everyday Musical Activities, Daily Listening and Everyday Musical Activities FAQ, My Daily Journal, Study Information,* and *Help and Assistance.* The first five pages mirrored the MBSH site in terms of structure, but differed in the text, audio recordings, and video used. References to mindfulness were replaced by information about classical music. Instead of the introductory video about mindfulness on the *What is Mindfulness?* page, the music site contained a brief introductory video about types of music and the potential benefits of listening to classical music. On the *Daily Classical Music Listening* page, two ten-minute classical music pieces replaced

the ten-minute mindfulness recordings (Beethoven's Piano Concerto No. 3 and No. 5). Consistent with previous research exploring the potential benefits of listening to classical music (e.g., Burns et al., 2002; Labbé et al., 2007), these pieces were selected based on their slow tempo markings. The final three pages were identical to the ones on the MBSH site.

Measures

All measures were completed at baseline, mid-intervention, and post-intervention with the exception of the plausibility question, which was completed at baseline only, the Engagement Questionnaire, which was completed at post-intervention only, and the Perceived Stress Scale (PSS), which was completed at baseline and post-intervention only. Cronbach's alphas used baseline data from all participants ($N = 214$).

Five-facet mindfulness questionnaire short form (FFMQ). The 24-item FFMQ (Bohlmeijer, ten Klooster, Fledderus, Veehof, & Baer, 2011) is a shortened version of the original 39-item measure (Baer, Smith, Hopkins, Kreitemeyer, & Toney, 2006). Both versions measure the general tendency to be mindful. The 24-item FFMQ is comprised of the same five facets as the original version: observing, describing, acting with awareness, non-judging of inner experience, and non-reactivity to inner experience. Total scale and subscale scores of the 24-item FFMQ were found to be highly correlated with the 39-item version. Facets of the 24-item FFMQ were also found to be as similarly sensitive to change as the original measure. Items are rated on a five-point Likert scale ranging from 1 (never or very rarely true) to 5 (very often or always true). Consistent with the design of this study, participants were asked to complete this measure based on their experiences in the past week, and only the total FFMQ score was used. The total FFMQ scores at each time point did not include items from the observing facet, in line with recommendations for excluding this facet from comparisons of total FFMQ scores before and after mindfulness interventions (e.g., Gu, Strauss, Crane et al., 2016). This recommendation is based on findings demonstrating differences in FFMQ factor structure before and after mindfulness training in the same sample; before mindfulness training, a four-factor hierarchical model (without the observing facet) best fit the data, but after mindfulness training, a five-factor hierarchical model (with all five facets) best fit data (Gu, Strauss, Crane et al., 2016).

Cronbach's alpha in this study for the total FFMQ score (excluding observing items) was .87.

Self-compassion scale short form (SCS). The 12-item SCS (Raes, Pommier, Neff, & Van Gucht, 2011) is a shortened version of the original 26-item version (Neff, 2003b). Confirmatory factor analysis supported a six-factor hierarchical structure of the measure with the same six factors as the original version: self-kindness, self-judgement, common humanity, isolation, mindfulness, and over-identification. However, internal consistency was variable for individual subscale scores and it is recommended that only the total SCS score is used. Items are rated on a five-point Likert scale from 1 (almost never) to 5 (almost always). Cronbach's alpha in this study for the total SCS score was .83.

Penn state worry questionnaire (PSWQ). The 16-item PSWQ (Meyer, Miller, Metzger, & Borkovec, 1990) is a widely used measure of trait worry. Each item is rated on a five-point Likert scale from 1 (not at all typical of me) to 5 (very typical of me). In the current study, participants completed this measure based on their experiences in the past week. Cronbach's alpha in this study for the total PSWQ score was .91.

Perceived stress scale (PSS). The 10-item PSS (Cohen, Kamarck, & Mermelstein, 1983) measures participants' perceptions of situations in their life as stressful. Items ask participants to rate how often they have thought or felt in a certain way during the last month using a five-point Likert scale ranging from 0 (never) to 4 (very often). Participants completed this measure based on their experiences in the past week. Cronbach's alpha in this study for the total PSS score was = .85.

Engagement questionnaire. The engagement questionnaire, designed by the research team, consisted of the following four questions: 1) "Over the past two weeks, how much time in total have you spent using the online site, not including time spent listening to the audio recordings?", 2) "Over the past two weeks, on how many days have you spent using the online site, not including time spent listening to the audio recordings?", 3) "Over the past two weeks, how much time in total have you spent listening to the audio recordings?", and 4) "Over the past two weeks, on how many days have you spent listening to the audio recordings?". Participants answered questions 1 and 3 by entering the number of minutes, and questions 2 and 4 by selecting a number from 0 to 14.

Table 4.

Baseline demographic characteristics of all participants across MBSH, classical music, and waitlist control groups

Variable	Total (N = 214)		MBSH (n = 83)		Music (n = 68)		Waitlist (n = 63)	
	M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Age	24.20	5.79	24.94	6.89	24.38	5.76	23.02	3.81
Gender	***N***	***%***	***N***	***%***	***N***	***%***	***N***	***%***
Male	58	27.10	18	21.69	19	27.94	21	33.33
Female	156	72.90	65	78.31	49	72.06	42	66.67
Occupation								
Student	187	87.38	67	80.72	59	86.76	61	96.83
Staff	27	12.62	16	19.28	9	13.24	2	3.17
Length of mindfulness practice								
No experience	117	54.67	43	51.81	36	52.94	38	60.32
Less than a year	62	28.97	29	34.94	19	27.94	14	22.22
1-5 years	31	14.49	11	13.25	11	16.18	9	14.29
Over 5 years	4	1.87	0	0	2	2.94	2	3.17

Frequency of mindfulness practice								
Not at all	120	56.07	43	51.81	37	54.41	40	63.49
Once a month or less	60	28.04	26	31.33	20	29.41	14	22.22
About once a week	30	14.02	13	15.66	9	13.24	8	12.70
Most days	4	1.87	1	1.20	2	2.94	1	1.59
Length of CM listening								
No experience	96	44.86	33	39.76	32	47.06	31	49.21
Less than a year	26	12.15	9	10.84	12	17.65	5	7.94
1-5 years	27	12.62	13	15.66	7	10.29	7	11.11
Over 5 years	65	30.37	28	33.73	17	25.00	20	31.75
Frequency of CM listening								
Not at all	95	44.39	32	38.55	33	48.53	30	47.62
Once a month or less	76	35.51	37	44.58	19	27.94	20	31.75
About once a week	35	16.36	11	13.25	14	20.59	10	15.87
Most days	8	3.74	3	3.61	2	2.94	3	4.76

Note. CM = classical music; MBSH = mindfulness-based self-help.

Plausibility question. Participants were asked about their perceived plausibility of the MBSH and classical music interventions using the following question: "On a scale from 1 (not at all) to 9 (very much), how much do you feel that mindfulness meditation/classical music will help your wellbeing?".

Data Analyses

Preliminary analyses. Pearson's chi-square test was conducted to examine between-group differences in the completion of questionnaires. Independent t-tests were conducted to investigate baseline differences between completers and non-completers of all three sets of questionnaires. Pearson's chi-square and one-way independent analysis of variance (ANOVA) were conducted to determine whether demographic variables differed between groups. Only participants who completed all three sets of questionnaires (baseline, mid-intervention, and post-intervention) were included in the main and mediation analyses.

Hypothesis 1: Main effects. A three (group: MBSH, music, and waitlist) by two (time: baseline and post-intervention) mixed ANOVA was conducted on stress scores. A significant interaction effect was followed up with an analysis of covariance (ANCOVA) test and planned simple contrasts, to examine between-group differences in post-intervention stress scores controlling for baseline stress scores.

Hypotheses 2 and 3: Mediation analyses. Mediation analyses used standardised residualised change scores for mediator and outcome variables. Standardised residuals were calculated using a linear regression model in which baseline scores predicted mid-intervention scores for mediators (mindfulness, self-compassion, and worry) and baseline scores predicted post-intervention scores for the outcome variable (perceived stress). As the independent variable (IV), group, is multicategorical (with 3 levels: MBSH, classical music, and waitlist) rather than dichotomous, three multicategorical IV mediation models were tested (Figure 4), one for each mediator, as recommended by Hayes and Preacher (2014). In each model, two IV contrasts were examined using MBSH as the reference category; MBSH versus music and MBSH versus waitlist control. Testing multiple mediation models was unsuitable given the theoretical and empirical overlap between mindfulness and worry (e.g., Feldman, Hayes, Kumar, Greeson, & Laurenceau, 2007; Verplanken & Fisher, 2013) and mindfulness and self-compassion (e.g., Baer, Lykins, & Peters, 2012); this

would test mindfulness as a mediator after controlling for worry and self-compassion as mediators (and vice versa), rather than the overall ability of mindfulness, self-compassion, and worry as mediators (Preacher & Hayes, 2008).

The three multicategorical IV models were tested using bias-corrected (BC) bootstrapping implemented in Mplus version 7.4 (Muthén & Muthén, 1998-2015). Point estimates of each indirect effect (*ab*) were calculated by averaging the *ab* product from 5,000 random samples of the original data. Indirect effects are significant if the upper and lower boundaries of the bootstrapped 95% BC confidence intervals (CIs) do not contain zero. Path coefficients were also calculated in Mplus for the effect of the IV on the mediator (path *a*), the effect of the mediator on the DV controlling for the IV (path *b*), and the direct effect of the IV on the DV controlling for the mediator (path *c'*). The path coefficients for the total effect of the IV on the DV *not* controlling for the mediator (path *c*) were calculated using linear regression in SPSS. Cohen's *d* effect sizes were calculated for between-group (MBSH versus music, MBSH versus waitlist) effects on mediator and outcome variables (paths *a* and *c*).

Results

Of the 214 participants randomised, 120 (56.07%) completed all three sets of measures and were included in the main and mediation analyses. Participant flow through the study is presented in Figure 5. Table 5 shows the demographic characteristics of the completer sample; chi-square tests and a one-way ANOVA showed that the only demographic variable which significantly differed across groups at baseline was occupation.

Completion rates were not found to significantly differ across groups ($\chi^2(2) = 1.93$, $p = .381$). Chi-square tests with Bonferroni-corrected alpha levels of .0083 (.05/6) showed no significant differences between completers and non-completers in gender, occupation, classical music experience (length of experience and frequency of listening to music), and mindfulness experience (length of experience and frequency of mindfulness practice). Independent t-tests with Bonferroni-corrected alpha levels of .01 (.05/5) demonstrated no significant differences between completers and non-completers in terms of age and baseline self-compassion, worry, and stress scores. However, completers were found to have significantly higher baseline mindfulness scores (*M* =

77.34, *SD* = 10.66) compared to non-completers (*M* = 73.57, *SD* = 8.91) (*p* = .005). Mean baseline, mid-intervention, and post-intervention scores for all measures across groups and time points are shown in Table 6.

Hypothesis 1: Main Effects

Mixed ANOVA showed a significant group by time interaction for perceived stress, $F(2, 117) = 3.35$, $p = .038$. Follow-up ANCOVA showed a significant effect of group on post-intervention stress scores controlling for baseline stress scores, $F(2, 116) = 3.77$, $p = .026$; post-intervention stress scores were significantly lower in the MBSH group compared to the waitlist control (contrast estimate = 3.35, $p = .014$) and in the MBSH compared to the music group (contrast estimate = 2.86, $p = .029$). These findings supported hypothesis 1.

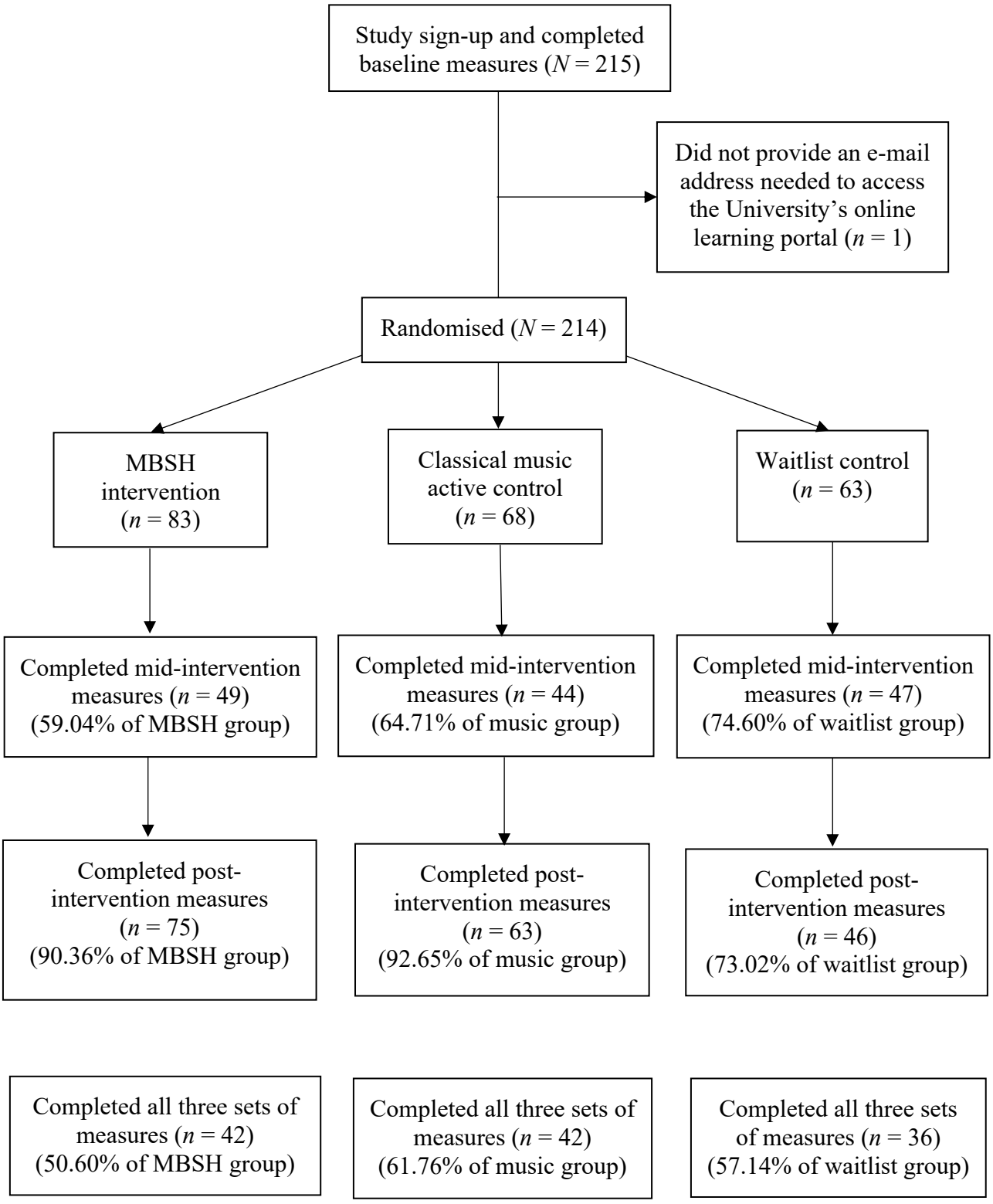


Figure 5. Participant flow through the study.

Table 5.

Baseline demographic characteristics of participants across MBSH, classical music, and waitlist control groups who completed all three sets of questionnaires

Variable	Total (n = 120)		MBSH (n = 42)		Music (n = 42)		Waitlist (n = 36)		Statistics[a]
	M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	
Age	24.66	6.40	25.64	8.10	24.33	6.23	23.89	3.85	$F(2, 117) = 0.81$, ns
Gender	***N***	***%***	***N***	***%***	***N***	***%***	***N***	***%***	$\chi^2(2) = 0.27$, ns
Male	36	30	12	28.57	12	28.57	12	33.33	
Female	84	70	30	71.43	30	71.43	24	66.67	
Occupation									$\chi^2(2) = 6.48$, $p = .04$
Student	109	90.83	35	83.33	38	90.48	36	100	
Staff	11	9.17	7	16.67	4	9.52	0	0	
Length of mindfulness practice									$\chi^2(6) = 6.90$, ns
No experience	70	58.33	23	54.76	25	59.52	22	61.11	
Less than a year	28	23.33	13	30.95	9	21.43	6	16.67	
1-5 years	20	16.67	6	14.29	8	19.05	6	16.67	
Over 5 years	2	1.67	0	0	0	0	2	5.56	

Frequency of mindfulness practice									$\chi^2(6) = 2.66$, ns
Not at all	71	59.17	23	54.76	25	59.52	23	63.89	
Once a month or less	27	22.50	9	21.43	10	23.81	8	22.22	
About once a week	20	16.67	9	21.43	7	16.67	4	11.11	
Most days	2	1.67	1	2.38	0	0	1	2.78	
Length of CM listening									$\chi^2(6) = 2.38$, ns
No experience	52	43.33	16	38.10	20	47.62	16	44.44	
Less than a year	8	6.67	2	4.76	4	9.52	2	5.56	
1-5 years	16	13.33	7	16.67	4	9.52	5	13.89	
Over 5 years	44	36.67	17	40.48	14	33.33	13	36.11	
Frequency of CM listening									$\chi^2(6) = 4.53$, ns
Not at all	50	41.67	15	35.71	19	45.24	16	44.44	
Once a month or less	46	38.33	21	50.00	14	33.33	11	30.56	
About once a week	20	16.67	5	11.90	8	19.05	7	19.44	
Most days	4	3.33	1	2.38	1	2.38	2	5.56	

Note. CM = classical music; MBSH = mindfulness-based self-help.

[a] Statistical test for between-group differences in demographic characteristics of participants at baseline. The only demographic variable which significantly differed across groups at baseline was occupation.

Table 6.

Mean total mindfulness, self-compassion, worry, and perceived stress scores across all conditions and time points in the completer sample (n = 120)

	MBSH (*n* = 42)			Music (*n* = 42)			Waitlist (*n* = 36)		
	Pre	Mid	Post	Pre	Mid	Post	Pre	Mid	Post
FFMQ[a]	62.48	67.10	67.45	63.26	64.07	65.90	64.11	64.89	64.75
	(10.05)	(10.82)	(11.02)	(10.00)	(10.27)	(9.96)	(11.18)	(11.80)	(13.36)
SCS	33.86	37.38	37.83	35.10	36.00	37.64	36.94	37.11	37.25
	(9.24)	(8.80)	(8.12)	(7.75)	(7.10)	(7.40)	(9.31)	(10.29)	(10.63)
PSWQ	53.79	47.26	47.00	51.19	50.55	49.48	51.78	51.33	49.47
	(12.29)	(11.80)	(10.82)	(11.98)	(11.32)	(11.17)	(15.43)	(15.21)	(17.02)
PSS	28.88	-	25.48	28.83	-	28.31	27.97	-	28.33
	(7.41)		(6.63)	(6.84)		(7.06)	(8.27)		(7.90)

Note. FFMQ = Five-Facet Mindfulness Questionnaire; MBSH = mindfulness-based self-help; PSS = Perceived Stress Scale; PSWQ = Penn State Worry Questionnaire; SCS = Self-Compassion Scale. Standard deviations are shown in parentheses.

[a] Mean total FFMQ scores do not include items from the observing facet.

Hypotheses 2 and 3: Mediation Analyses

Table 7 presents the results of the mediation analyses for the three multicategorical IV mediation models that tested whether standardised residualised change scores in mindfulness, self-compassion, and worry mediated the relationship between group (MBSH vs. waitlist, MBSH vs. music) and standardised residualised change scores in perceived stress. Figures 6, 7, and 8 present path diagrams of these models.

Hypothesis 2. Participation in MBSH versus waitlist control on change in perceived stress was hypothesised to be significantly mediated by improvements in mindfulness, self-compassion, and worry. Change in mindfulness was found to significantly mediate the effects of participating in MBSH versus waitlist on changes in stress (model 1, contrast 2), as indicated by the bootstrapped 95% BC CI which did not cross zero (Table 7). All path coefficients apart from the one corresponding to path a_2 were significant. This shows that there was a significant and moderate effect of group (MBSH versus waitlist) on change in stress (path c_2) in favour of the MBSH group (d = 0.62) and that change in mindfulness was a significant predictor of change in stress (path b_1). Although the effect of group on change in mindfulness just failed to reach statistical significance (path a_2; p =.060), there was a moderate effect size (d = 0.43) in favour of the MBSH group.

Change in self-compassion was found to be a significant mediator of the effects of MBSH versus waitlist on change in stress (model 2, contrast 2), as indicated by the bootstrapped 95% BC CI which did not include zero. All paths in this model were significant. In addition to the significant, moderate effect of group (MBSH versus waitlist) on changes in stress (path c_2; d = 0.62), there was a significant, moderate effect of group on change in self-compassion (path a_4) in favour of the MBSH group (d = 0.45) and change in self-compassion was a found to be a significant predictor of changes in stress (path b_2).

The effects of MBSH versus waitlist on change in stress was significantly mediated by changes in worry (model 3, contrast 2), as indicated by the bootstrapped 95% BC CI which did not include zero. All paths in this model were significant. In addition to the significant, moderate effect of group (MBSH versus waitlist) on change in stress (path c_2; d = 0.62), there was a significant, large effect of group on change in worry (path a_6), in favour of the MBSH group (d = 0.77), and change in worry was a

found to be a significant predictor of change in stress (path b_3). Taken together, these results supported hypothesis 2.

Hypothesis 3. Participation in the MBSH versus classical music condition on change in perceived stress was predicted to be significantly mediated by improvements in mindfulness, self-compassion, and worry. Change in mindfulness was found to significantly mediate the effect of participating in MBSH versus music on changes in perceived stress (model 1, contrast 1), as indicated by bootstrapped 95% BC CIs which did not contain zero. All regression coefficients apart from the one corresponding to path c'_1 were significant. In addition to change in mindfulness being a significant predictor of change in stress (path b_1), there was a significant, moderate effect of group (MBSH versus classical music) on change in stress (path c_1), in favour of the MBSH group ($d = 0.46$), and significant, moderate effect of group on change in mindfulness (path a_1) in favour of the MBSH group ($d = 0.46$).

Change in self-compassion was found to significantly mediate the effects of MBSH versus music on change in stress (model 2, contrast 1), as indicated by bootstrapped CIs which did not cross zero. All paths in this model apart from the one corresponding to path c'_3 were significant. In addition to change in self-compassion being a significant predictor of change in stress (path b_2) and a significant, moderate effect of group (MBSH versus classical music) on change in stress (path c_1; $d = 0.46$), there was a significant, moderate effect of group on change in self-compassion (path a_3) in favour of the MBSH group ($d = 0.43$).

Change in worry was found to significantly mediate the effects of MBSH versus music on change in stress (model 2, contrast 1), as indicated by bootstrapped CIs which did not include zero. All paths in this model apart from the one corresponding to path c'_5 were significant. In addition to change in worry being a significant predictor of change in stress (path b_3) and a significant, moderate effect of group (MBSH versus classical music) on change in stress (path c_1; $d = 0.46$), there was a significant, moderate-large effect of group on change in worry (path a_5) in favour of the MBSH group ($d = 0.70$). Taken together, these findings supported hypothesis 3.

Table 7.

Unstandardised regression coefficients, their standard errors (SEs) and significance values, and bootstrapped unstandardised point estimates, and their SEs and 95% bias-corrected confidence intervals, for the three multicategorical independent variable mediation models

Model	*B*	*SE*	*t*	*p*	Point estimate (*SE*) [95% BC CIs][a]
Model 1. With **mindfulness** as the mediator.					
Contrast 1. Group: MBSH vs. music					0.13 (0.08) [0.01, 0.15]
a_1 path: group -> FFMQ change	-0.46	0.22	-2.06	.040	
b_1 path: FFMQ change -> PSS change	-0.28	0.09	-3.15	.002	
c_1 path: group -> PSS change	0.47	0.23	2.10	.039	
c'_1 path: group -> PSS change (direct effect)	0.35	0.22	1.58	.114	
Contrast 2. Group: MBSH vs. waitlist control					0.12 (0.08) [0.002, 0.15]
a_2 path: group -> FFMQ change	-0.43	0.23	-1.88	.060	
b_1 path: FFMQ change -> PSS change	-0.28	0.09	-3.15	.002	
c_2 path: group -> PSS change	0.28	0.10	2.72	.008	
c'_2 path: group -> PSS change (direct effect)	0.43	0.19	2.28	.023	
Model 2. With **self-compassion** as the mediator.					

Contrast 1. Group: MBSH vs. music					0.09 (0.06) [0.01, 0.25]
a_3 path: group -> SCS change	-0.42	0.21	-1.99	.047	
b_2 path: SCS change -> PSS change	-0.22	0.09	-2.54	.011	
c_1 path: group -> PSS change	0.47	0.23	2.10	.039	
c'_3 path: group -> PSS change (direct effect)	0.38	0.23	1.67	.094	
Contrast 2. Group: MBSH vs. waitlist control					0.10 (0.07) [0.01, 0.28]
a_4 path: group -> SCS change	-0.47	0.24	-2.00	.045	
b_2 path: SCS change -> PSS change	-0.22	0.09	-2.54	.011	
c_2 path: group -> PSS change	0.28	0.10	2.72	.008	
c'_4 path: group -> PSS change (direct effect)	0.45	0.20	2.30	.021	
Model 3. With **worry** as the mediator.					
Contrast 1. Group: MBSH vs. music					0.16 (0.07) [0.05, 0.33]
a_5 path: group -> PSWQ change	0.66	0.21	3.23	.001	
b_3 path: PSWQ change -> PSS change	0.24	0.09	2.58	.010	
c_1 path: group -> PSS change	0.47	0.23	2.10	.039	
c'_5 path: group -> PSS change (direct effect)	0.32	0.23	1.38	.166	

Contrast 2. Group: MBSH vs. waitlist control					0.17 (0.08) [0.05, 0.37]
a_6 path: group -> PSWQ change	0.71	0.21	3.43	.001	
b_3 path: PSWQ change -> PSS change	0.24	0.09	2.58	.010	
c_2 path: group -> PSS change	0.28	0.10	2.72	.008	
c'_6 path: group -> PSS change (direct effect)	0.39	0.19	2.00	.045	

Note. Standardised residualised change scores were used for all mediator and outcome variables. BC CIs = bias-corrected confidence intervals; MBSH = mindfulness-based self-help; PSS = Perceived Stress Scale; PSWQ = Penn State Worry Questionnaire; SCS = Self-Compassion Scale.

[a] Bootstrapped 95% BC CIs for the *ab* (indirect) effect; a significant indirect effect is indicated where these do not cross zero ($p < .05$).

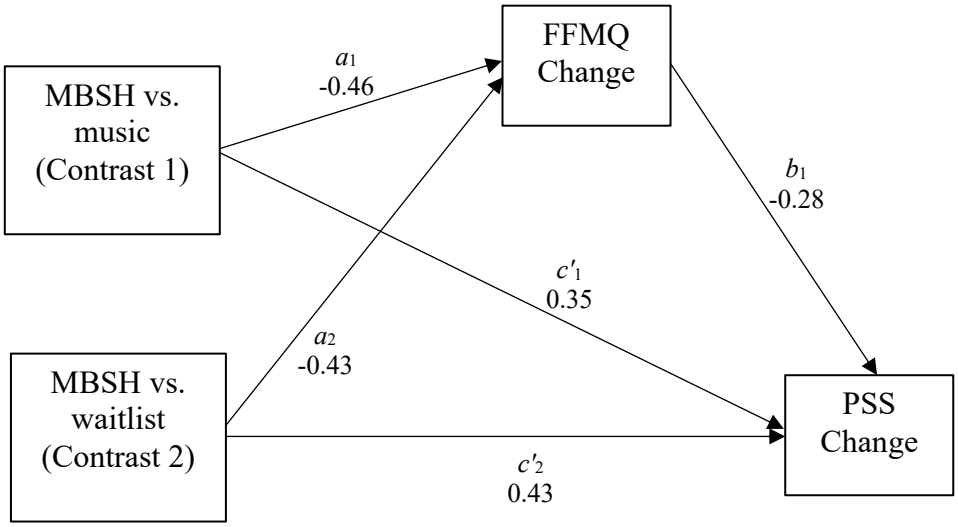


Figure 6. Path diagram depicting Model 1, testing whether changes in mindfulness (FFMQ Change) mediate the effects of mindfulness-based self-help (MBSH) versus music (Contrast 1) or MBSH versus waitlist control (Contrast 2) on improvements in perceived stress (PSS Change). Unstandardised path coefficients are displayed. Change refers to standardised residualised change scores.

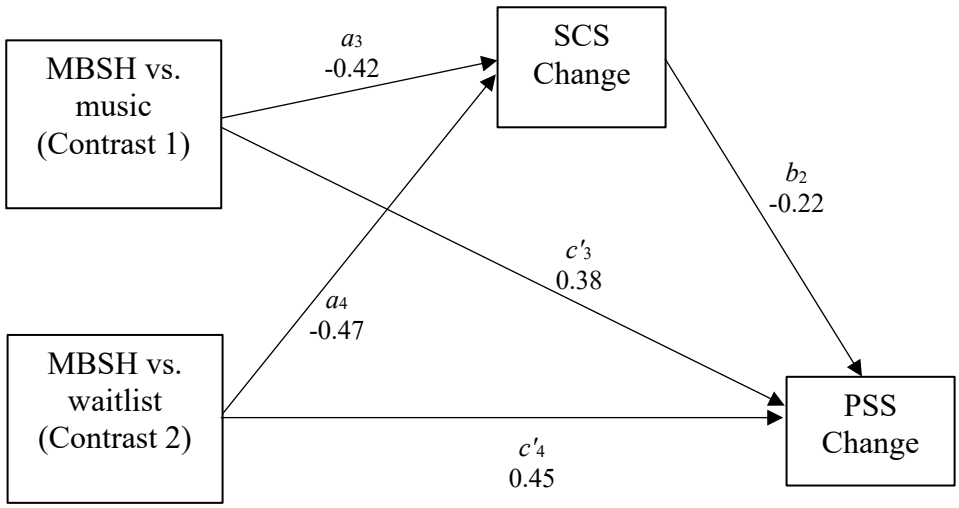


Figure 7. Path diagram depicting Model 2, testing whether changes in self-compassion (SCS Change) mediate the effects of mindfulness-based self-help (MBSH) versus music (Contrast 1) or MBSH versus waitlist control (Contrast 2) on improvements in perceived stress (PSS Change). Unstandardised path coefficients are displayed. Change refers to standardised residualised change scores.

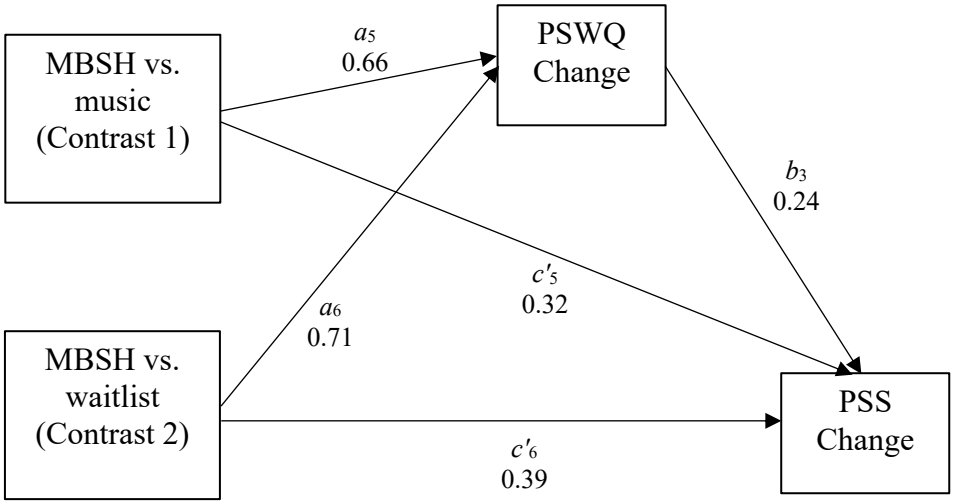


Figure 8. Path diagram depicting Model 3, testing whether changes in worry (PSWQ Change) mediate the effects of mindfulness-based self-help (MBSH) versus music (Contrast 1) or MBSH versus waitlist control (Contrast 2) on improvements in perceived stress (PSS Change). Unstandardised path coefficients are displayed. Change refers to standardised residualised change scores.

Intervention Engagement

There were no significant differences between the MBSH and music conditions on any of the four engagement indices; time spent browsing the site ($t(82) = 0.35$, $p = .724$), number of days spent browsing the site ($t(75.18) = 1.46$, $p = .148$), time spent listening to audio recordings ($t(82) = 1.07$, $p = .286$), and number of days spent listening to audio recordings ($t(82) = 0.34$, $p = .732$).

Participants in the MBSH condition reported spending an average of 72.98 minutes ($SD = 63.93$) and 5.12 days ($SD = 3.66$) over the two weeks browsing the online site, not including time spent listening to audio recordings. They reported spending on average 99.43 minutes ($SD = 83.64$) and 7.52 days ($SD = 4.52$) over the two weeks listening to audio recordings on the site.

Participants in the classical music condition reported spending an average of 79.05 minutes ($SD = 90.86$) and 4.10 days ($SD = 2.69$) over the two weeks browsing the online site, not including time spent listening to audio recordings. They reported spending on average 121.33 minutes ($SD = 102.47$) and 7.21 days ($SD = 3.68$) over the two weeks listening to audio recordings on the site.

Intervention Plausibility

For the completer sample, there was no significant difference in perceived plausibility of received intervention between participants randomised to the MBSH intervention ($M = 5.74$, $SD = 2.04$) and participants randomised to the music intervention ($M = 5.02$, $SD = 1.99$), $t(82) = 1.62$, $p = .108$.

Discussion

This experimental study examined the effects of a two-week online MBSH intervention on perceived stress compared to matched classical music and inactive waitlist control groups. This study also tested whether improvements in three theoretically and/or empirically supported mechanisms of MBIs (mindfulness, self-compassion, and worry) mediate the effects of MBSH compared to both control conditions on changes in stress.

Consistent with hypothesis 1, MBSH was found to significantly reduce stress at post-intervention compared to both the matched classical music condition and the

waitlist condition. In addition, MBSH had significant, moderate-to-large effects on improvements in not only stress, but also mindfulness, self-compassion, and worry over the course of the intervention (indicated by significant coefficients for paths *c* in the mediation analyses), in comparison to both control conditions. Findings support and extend the modest body of evidence for the effectiveness of MBSH interventions (Cavanagh et al., 2014). Given that listening to classical music was rated by participants as equally plausible and engaging as MBSH, the finding that MBSH had significant effects on stress compared to the music intervention suggests that MBSH may be a particularly effective way of managing stress in a non-clinical population.

Both mediational hypotheses were supported; improvements in mindfulness, self-compassion, and worry significantly mediated the effects of MBSH on changes in stress in comparison to a waitlist control group and a matched, equally plausible and engaging classical music intervention. This suggests that the mediating effects of mindfulness, self-compassion, and worry on stress outcomes are specific to learning mindfulness and not general features of any plausible self-help intervention. The inclusion of a well-matched control group allows stronger inferences to be made regarding specificity of effects to MBSH.

These mediation findings suggest that there is an overlap between the mechanisms underlying MBIs and MBSH interventions; current findings are consistent with findings identifying mindfulness and worry as two of the most empirically supported mediators of the effects of MBIs on mental health outcomes (Gu, Strauss, Bond, & Cavanagh, 2016). Findings also support the theoretical literature on MBIs; evidence for mindfulness as a mediator supports the theoretical premise of MBIs such as MBCT and MBSR, that cultivating mindfulness improves mental health outcomes (Kabat-Zinn, 1982; Segal et al., 2002, 2013). Findings are also consistent with the notion that self-compassion is embedded in mindfulness practice and crucial to the change process (Feldman & Kuyken, 2011). Evidence for worry as a mediator supports MBCT theory that improved mental health outcomes, in particular reduced depressive relapse, are a result of decreasing maladaptive styles of thinking (Segal et al., 2002, 2013).

Strengths and Limitations

This study addressed an important omission in the MBI literature investigating effectiveness and mechanisms. The use of MBSH allowed us to examine the specific effects associated with learning mindfulness, by removing many of the non-specific factors found in group-based MBIs (e.g., facilitator support, group process) and MBIs which incorporate cognitive behavioural therapy elements (e.g., MBCT) and stress reduction strategies (e.g., MBSR). By comparing MBSH with a well-matched non-mindfulness intervention, reported to be comparably plausible and engaging, this study also controlled for additional non-specific factors (e.g., expectation of benefit). This allows for stronger conclusions regarding the specificity of effects of learning mindfulness (i.e., mediating effects are not simply features of any plausible self-help intervention).

However, study attrition was high, with only 56.07% of participants completing measures at all three time points. Although not unusual in studies of online MBSH (e.g., Cavanagh et al., 2013), relatively low rates of study completion (i.e., completing measures at all three time points) can be largely attributed to participants who did not complete measures mid-intervention. Low rates of measure completion at mid-intervention may be because measures needed to be completed within a short time frame and as a result, only one reminder email was sent. Future studies incorporating assessment points during a brief intervention could inform participants of upcoming assessments (e.g., 24 hours prior to sending questionnaires) in addition to sending multiple reminder e-mails.

Although the current study contributes to our understanding of how learning mindfulness might reduce stress, a more robust test of the underlying mechanisms would involve determining the temporal order of mediator and outcome variables, by assessing these variables during the intervention (e.g., at mid-intervention) and testing whether change in mediators predates change in outcomes (Kazdin, 2007). In the current study, mindfulness, self-compassion, and worry were found to be significant mediators of MBSH's specific effects on stress, but we cannot infer causal direction from our data, because we do not know if changes in the proposed mediators improved prior to or following changes in stress. Future research should endeavour to include assessment of both mediator and outcome variables during intervention in order to more closely examine temporal ordering and make more definitive conclusions regarding the causal mechanisms of MBIs.

Future studies should also follow up participants to determine whether improvements in outcomes gained from participating in MBSH are maintained beyond the intervention period. Further research may additionally benefit from using objective measures of engagement (e.g., number of times web pages and audio recordings were accessed) and including non-self-report measures of mediator and outcome variables.

Current findings show that online MBSH interventions may be promising alternatives to MBIs in non-clinical settings and in situations where resources are limited. However, a cautious approach to applying these findings to clinical practice should be taken. There is currently very limited evidence that MBSH can be offered safely and effectively in clinical settings (although see Dimidjian et al., 2014 for a recent promising example). As such, we do not recommend at present that MBSH, offered without support from a qualified mindfulness teacher, should be routinely offered in clinical settings. Further research is needed to test the safety, acceptability, and effectiveness of MBSH in clinical populations and examine whether similar mechanisms are involved. MBSH in clinical settings may also require the addition of clinician guidance and support in order to maximise engagement and improve outcomes, as has been found for self-help CBT in these settings (cf. Gilbody et al., 2015).

Previous research examining the effects of MBIs and mechanisms of change has been hampered by failure to control for non-specific effects such as group process, facilitator support, and expectation of benefit. This means that it remained possible that the benefits of MBIs could have been attributable, perhaps entirely, to these non-specific factors. Findings from the current study however suggest that mindfulness-specific factors do contribute, at least in part, to beneficial outcomes and change processes. Compared to a waitlist control group and a well-matched non-mindfulness condition, designed to control for non-specific effects, MBSH led to significant improvements in stress. Moreover, changes in mindfulness, self-compassion, and worry significantly mediated the effects of MBSH versus both control conditions on stress reduction. Future research on mechanisms should include multiple assessments of mediator and outcome variables during intervention to draw robust conclusions regarding direction of causality. These findings demonstrate that learning mindfulness *per se* may confer specific benefits to mental health, through improving mindfulness and self-compassion and reducing worry.

Chapter 4:
Examining the Factor Structure of the 39-Item and 15-Item Versions of the Five Facet Mindfulness Questionnaire Before and After Mindfulness-Based Cognitive Therapy for People with Recurrent Depression

Chapter 4 is published in *Psychological Assessment* as:

Gu, J., Strauss, C., Crane, C., Barnhofer, T., Karl, A., Cavanagh, K., & Kuyken, W. (2016). Examining the Factor Structure of the 39-Item and 15-Item Versions of the Five Facet Mindfulness Questionnaire Before and After Mindfulness-Based Cognitive Therapy for People with Recurrent Depression. *Psychological Assessment, 28*(7), 791.

Author Contributions

Design:	WK, CC, CS, and JG
Data Collection:	N/A*
Data Analysis:	JG
Initial Draft:	JG
Reviewing and Editing:	All authors

*Secondary data analysis of data collected from two randomised controlled trials.

Abstract

Research into the effectiveness and mechanisms of mindfulness-based interventions (MBIs) requires reliable and valid measures of mindfulness. The 39-item Five-Facet Mindfulness Questionnaire (FFMQ-39) is a measure of mindfulness commonly used to assess change before and after MBIs. However, the stability and invariance of its factor structure have not yet been tested before and after an MBI; pre to post comparisons may not be valid if the structure changes over this period. Our primary aim was to examine the factor structure of the FFMQ-39 before and after Mindfulness-Based Cognitive Therapy (MBCT) in adults with recurrent depression in remission using confirmatory factor analysis (CFA). Additionally, we examined whether the factor structure of the 15-item version (FFMQ-15) was consistent with that of the FFMQ-39, and whether it was stable over MBCT. Our secondary aim was to assess the general psychometric properties of both versions. CFAs showed that pre-MBCT, a four-factor hierarchical model (excluding the 'observing' facet) best fit the FFMQ-39 and FFMQ-15 data, whereas post-MBCT, a five-factor hierarchical model best fit the data for both versions. Configural invariance across the time points was not supported for both versions. Internal consistency and sensitivity to change were adequate for both versions. The two FFMQ versions did not differ significantly from each other in terms of convergent validity. Researchers should consider excluding the 'observing' subscale from comparisons of total scale/subscale scores before and after mindfulness interventions. Current findings support the use of the FFMQ-15 as an alternative measure in research where briefer forms are needed.

Keywords: mindfulness, MBCT, FFMQ, confirmatory factor analysis, configural invariance

Introduction

Mindfulness is commonly defined as "paying attention in a particular way; on purpose, in the present moment, and non-judgmentally" (Kabat-Zinn, 1994, p. 4). It involves being aware and accepting of internal and external moment-to-moment experience and relating to thoughts and emotions in a decentred manner as 'mental events', rather than accurate reflections of the self and reality. Mindfulness is regarded as a universal human capacity that can enhance wellbeing (e.g., Ludwig & Kabat-Zinn, 2008). The secular practice of mindfulness has been integrated into various clinical interventions, with a view to increasing mindfulness and, as a consequence, improving mental health and wellbeing. The two most extensively applied and assessed mindfulness-based interventions (MBIs) are Mindfulness-Based Stress Reduction (MBSR; Kabat-Zinn, 1982) and Mindfulness-Based Cognitive Therapy (MBCT; Segal, Williams, & Teasdale, 2002, 2013). Both MBSR and MBCT are eight-session group-based programmes in which participants engage in formal and informal mindfulness meditation practices during sessions and at home. MBSR was developed to alleviate distress, pain, stress, and anxiety in people with chronic physical health problems through the cultivation of mindfulness. More recently, MBCT was designed for people with recurrent major depressive disorder (MDD) in remission as a relapse prevention intervention. MBCT is theorised to decrease depressive recurrence by enhancing mindful awareness of and disengagement from dysphoria-triggered repetitive negative thinking (e.g., rumination) about one's depressive symptoms (Segal et al., 2002, 2013).

Meta-analytic reviews have found MBCT and MBSR to be effective in improving a range of outcomes in clinical and non-clinical samples including stress, depression, depressive relapse, and anxiety (e.g., Chiesa & Serretti, 2009; Hofmann, Sawyer, Witt, & Oh, 2010; Piet & Hougaard, 2011; Strauss, Cavanagh, Oliver, & Pettman, 2014). As evidence for the effectiveness of MBIs is accumulating and these interventions are being adapted to target a broad range of problems, there is an increasing need for investigations of the mechanisms of change (see Gu, Strauss, Bond, & Cavanagh, 2015 for a review). At the very basis of this endeavour is whether MBIs work through their purported mechanisms of action (e.g., by increasing mindfulness) and which aspects of the construct of mindfulness are being affected by the training.

Psychological outcomes and processes from effectiveness and mechanism studies have been predominantly measured using self-report questionnaires, due to their

cost-effectiveness and standardised, easy-to-administer format. Among the available self-report measures of mindfulness, the Five-Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006) is a widely used measure that aims to capture the key underlying dimensions of mindfulness (Sauer et al., 2013).

The Five-Facet Mindfulness Questionnaire

The FFMQ is a 39-item (FFMQ-39) self-report measure of the dispositional tendency to be mindful in daily life. The questionnaire derived from an exploratory factor analysis (EFA) (Baer et al., 2006) of items from five independently developed self-report mindfulness scales: the (1) Mindfulness Attention Awareness Scale (Brown & Ryan, 2003), (2) Freiburg Mindfulness Inventory (Walach, Buchheld, Buttenmuller, Kleinknecht, & Schmidt, 2006), (3) Cognitive Affective Mindfulness Scale (Hayes & Feldman, 2004), (4) Southampton Mindfulness Questionnaire (Chadwick, Hember, Symes, Peters, Kuipers, & Dagnan, 2008), and (5) Kentucky Inventory of Mindfulness Skills (Baer, Smith, & Allen, 2004). Baer et al.'s (2006) findings showed that mindfulness can be conceptualised as a multifaceted construct consisting of five related dimensions: observing, describing, acting with awareness, non-judging of inner experience, and non-reactivity to inner experience. Observing refers to attending or noticing internal and external experiences (e.g., sounds, emotions, thoughts, bodily sensations, smells). Describing includes the ability to express in words one's experiences. Acting with awareness involves attending to one's present moment activity, rather than being on 'autopilot', or behaving automatically, while attention is focused elsewhere. Non-judging of inner experience involves accepting and not evaluating thoughts and emotions (e.g., as 'good' or 'bad'). Finally, non-reactivity to inner experience refers to the ability to detach from thoughts and emotions, allowing them to come and go without getting involved or carried away by them. The 39 items of the FFMQ are rated on a five-point Likert scale, ranging from 1 (never or very rarely true) to 5 (very often or always true). In addition to considering scores on the five subscales individually, facet scores can be combined to produce an overall mindfulness score.

Analyses of the psychometric properties of the FFMQ-39 have generally demonstrated this measure to have satisfactory convergent and discriminant validity, internal consistency, interpretability in distinguishing between participant subgroups,

and incremental validity in predicting psychological symptoms and wellbeing across samples of regular meditators and non-meditators (students, general community sample, adults with heterogeneous mood and anxiety disorders, adults with moderate depression or anxiety symptoms) (e.g., Baer et al., 2006; Baer et al., 2008; Bohlmeijer, ten Klooster, Fledderus, Veehof, & Baer, 2011; Christopher, Neuser, Michael, & Baitmangalkar, 2012; Curtiss & Klemanski, 2014a, 2014b). The FFMQ-39 has also been shown to have good sensitivity to change; researchers have found moderate-to-large, and statistically significant, increases in all five facets before and after MBSR in a sample of adults with chronic pain and heterogeneous mood and anxiety disorders (Carmody & Baer, 2008) and before and after a nine-week therapeutic intervention based on mindfulness in a sample of adults with mild to moderate symptoms of depression or anxiety (Bohlmeijer et al., 2011).

Although the psychometric properties of the FFMQ-39 have been supported, findings from a series of confirmatory factor analysis (CFA) studies question the validity of its five-factor structure and the inclusion of all five subscales in MBI research. The five-factor structure emerging in the development of the FFMQ-39 using EFA (Baer et al., 2006) has been confirmed in meditator samples, in which a five-factor hierarchical model (with the five related factors subsumed under an overarching mindfulness construct) provided the optimal fit for the data (Baer et al., 2008; M. J. Williams, Dalgleish, Karl, & Kuyken, 2014). However, for non-meditator samples (general community sample, students, adults with recurrent MDD in remission, adults with heterogeneous mood and anxiety disorders), a four-factor hierarchical model (with all facets minus observing loading onto an overall mindfulness factor) best fit the data (Baer et al., 2006, 2008; Curtiss & Klemanski, 2014a; M. J. Williams et al., 2014). Poor fit of the five-factor hierarchical model in non-meditator samples can be attributed to the non-significant relationships found between observing and non-judging (Baer et al., 2006; Bohlmeijer et al., 2011; Curtiss & Klemanski, 2014a), and observing and acting with awareness (Curtiss & Klemanski, 2014a).

A possible explanation for these differing factor structures across meditators and non-meditators is that the qualities of observing may differ in meditators and non-meditators, such that increased meditation practice strengthens the relationships between observing and the other mindfulness facets (Baer et al., 2008). For non-meditators, observing items (e.g., “When I’m walking, I deliberately notice the sensations of my body moving”) may be equally likely to reflect neutral attention, or

indeed even maladaptive, biased, and pathological forms of attention (e.g., anxious monitoring, hypervigilance to threat), rather than attention characterised by the curious, accepting, and purposeful quality cultivated through mindfulness meditation practice. Therefore, people with little or no mindfulness experience may report how much they tend to observe, but the way in which they notice may or may not be related to mindful qualities assessed by the other facets, resulting in the emergence of a four-factor hierarchical solution. By contrast, people with meditation experience may respond to observing items in a way which is more consistent with the other four facets and with a mindful disposition, thus resulting in the emergence of a five-factor hierarchical solution. In support of this explanation, findings show that in non-meditators (student, community, and highly educated samples), observing was the only facet which was positively correlated with psychological symptoms; in meditators, all facets negatively correlated with psychological symptoms (Baer et al., 2008).

To reduce participant burden in research trials, which include multiple measures and/or measures administered on multiple occasions, short versions of the original FFMQ-39 have been developed. One such version is a 24-item FFMQ, which has been shown to replicate the five-factor structure of the original measure, to be highly correlated with the original version, and to be sensitive to change (Bohlmeijer et al., 2011). More recently, a 15-item version (FFMQ-15) has been developed, which includes three items for each of the five facets (Baer, Carmody, & Hunsinger, 2012). These were selected based on the factor loadings for each subscale of the FFMQ-39 (Baer et al., 2006) and to maintain the breadth of content for each facet. However, the factor structure, correlation with the FFMQ-39, convergent validity, and sensitivity to change of the FFMQ-15 has yet to be validated.

Findings that highlight different factor structures for the FFMQ-39 in meditators and non-meditators (Baer et al., 2006, 2008; Curtiss & Klemanski, 2014a; M. J. Williams et al., 2014) have implications for studies using this measure to compare levels of mindfulness across these two samples or to evaluate the effectiveness of MBIs in samples with no previous meditation experience. Although the factor structure of the FFMQ-39 has been tested in a number of samples (e.g., meditators, students, general community sample, adults with recurrent MDD in remission, adults with heterogeneous mood and anxiety disorders), no known studies have yet directly examined the stability of the factor structure before and after mindfulness training (e.g., through MBCT) in a single sample. Previous studies suggest that it is meditation status that results in

differential factor structures emerging for the FFMQ-39; a study evaluating the measure's factor structure before and after an MBI in the same sample would provide a stronger test of whether mindfulness meditation experience changes the factor structure of the FFMQ-39.

The Present Study

The primary aim of this study was to examine the stability of the factor structure of the FFMQ-39 before and after MBCT using CFA. As the FFMQ-15 has not yet been validated, we also examined whether its factor structure was consistent with that of the original version, and whether the factor structure of the FFMQ-15 was invariant over a period in which people were learning mindfulness through MBCT. Data from two trials evaluating MBCT for adults with MDD in remission were used, meaning that the people contributing data were representative of the population for whom MBCT was originally designed. The secondary aim of this study was to assess the general psychometric properties of the FFMQ-39 and FFMQ-15. Each facet's sensitivity to change over the course of MBCT was examined. Convergent validity of the FFMQ-39 and FFMQ-15 were also tested by correlating the facets with theoretically related constructs before and after MBCT, specifically, measures of depression and negative rumination. Significant moderate negative correlations were expected between rumination and depression and the facets describing, acting with awareness, non-judging, and non-reactivity of both versions of the FFMQ. Given that research into how meditation experience might alter the way in which people observe is still emerging, no hypotheses were made regarding the correlations between the observing facet of the FFMQ-39 and FFMQ-15 and rumination and depression.

Method

Participants and Design

The sample consisted of participants from two trials which examined the effectiveness of MBCT compared to control conditions at reducing relapse into depression for people with recurrent MDD in remission (Preventing depressive relapse in NHS settings through MBCT [PREVENT] trial; Kuyken et al., 2015; and Staying Well After Depression [SWAD] trial; J. M. G. Williams et al., 2014). M. J. Williams et

al.'s (2014) CFA study also used data from PREVENT to examine the factor structure of the FFMQ-39 at baseline. However, our study used an extended sample and differed from theirs in the research questions tested; M. J. Williams et al. compared the FFMQ factor structure across independent samples of meditators and non-meditators at one time point, whereas we examined the stability of the FFMQ structure before and after MBCT in a single sample.

Both PREVENT and SWAD were multicentre trials, with PREVENT recruiting from general practices in rural and urban settings in the UK and SWAD recruiting from the community, primary care, and mental health clinics in the regions of Oxford, England, and Bangor, North Wales. Inclusion criteria for both trials included: (1) a diagnosis of recurrent MDD in full or partial remission according to the *Diagnostic and Statistical Manual of Mental Disorders - IV* (American Psychiatric Association, 1994), (2) three or more previous depressive episodes, and (3) being 18 years or older. Exclusion criteria from both trials included: having (1) a current major depressive episode, (2) a comorbid diagnosis of current substance misuse, organic brain damage, current or past psychosis, current or past bipolar disorder, persistent antisocial behaviour, or persistent self-harm requiring clinical management or therapy, and (3) formal concurrent psychotherapy. Only data from participants in the MBCT arm of both trials, who completed all FFMQ items both before and after MBCT and who took part in at least four out of eight sessions of MBCT (i.e., who were deemed therapy completers; Teasdale et al., 2000), were used in this study.

The total number of participants who fit the criteria was 238 (74.38% of the total number of participants randomised to MBCT in PREVENT and SWAD), 154 participants from the PREVENT trial and 84 participants from the SWAD trial. Of the 238 participants, 69 (29%) were men and 169 (71%) were women. The mean age of the sample was 49.18 years (SD = 12.01, range 23-78). Most (97.5%) of the sample were white. In terms of educational qualifications, 13 (5.55%) had no qualifications, 34 (14.3%) had some General Certificate of Secondary Education/O-Levels, 71 (29.8%) had some A Levels or comparable vocational qualifications, 57 (23.9%) had a bachelor's degree, 24 (10.1%) had a master's degree, and 30 (21.6%) had a doctoral degree or professional qualification. Three participants had other qualifications and data on education were missing for six.

Measures

Five-facet mindfulness questionnaire (FFMQ). The 39-item FFMQ (Baer et al., 2006) measures the trait-like tendency to be mindful in daily life. It is comprised of the following five related facets: observing, describing, acting with awareness, non-judging, and non-reactivity. Sample items include: "I notice the smells and aromas of things" (observing), "I'm good at finding words to describe my feelings" (describing), "I find myself doing things without paying attention" (acting with awareness), "I disapprove of myself when I have illogical ideas" (non-judging), and "When I have distressing thoughts or images, I don't let myself be carried away by them" (non-reactivity). Items are rated on a five-point Likert scale ranging from 1 (never or very rarely true) to 5 (very often or always true). Facet scores range from 8 to 40, with the exception of the non-reactivity facet, which ranges from 7 to 35. The 15-item FFMQ (Baer et al., 2012) includes the following items of the FFMQ-39 for each of the five facets: Items 6, 11, and 15 for observing, items 2, 16, and 27 for describing, items 8, 34, and 38 for acting with awareness, items 10, 14, and 30 for non-judging, and items 19, 29, and 33 for non-reactivity. These items were selected by Baer et al. (2012) based on their factor loadings and to maintain the range of content for each facet. The FFMQ-15 is measured using the same scale as the FFMQ-39 and its facet scores range from 3 to 15. In the current study only the FFMQ-39 was administered to participants; FFMQ-15 data were extracted from the 39-item version. Cronbach's alphas for facets from both versions of the measure are displayed in Table 8.

Beck depression inventory II (BDI-II). The 21-item BDI-II (Beck, Steer, & Brown, 1996) is widely used to assess the severity of depressive symptomatology. Each item is a list of four statements about a symptom of depression, arranged in order of severity. Items are rated on a 4-point scale ranging from 0 (not at all) to 3 (extreme form of each symptom), which corresponds to each statement. Items are summed to give a single total score which ranges from 0 to 63; a score of 0-13 is considered to reflect minimal depression, 14-19 mild depression, 20-29 moderate depression, and 30-63 severe depression. Cronbach's alpha was .93 for this sample at both baseline and post-MBCT.

Cambridge-Exeter ruminative thinking scale (CERTS). The CERTS (Barnard, Watkins, Mackintosh, & Nimmo-Smith, 2007) is a transdiagnostic tool for assessing multiple aspects of rumination. The measure consists of three parts, each with several subscales. The first scale measures patterns of ruminative thinking across

multiple contexts (e.g., anxious, happy, sad) and consists of two subscales: rumination in response to negative mood and negative exigencies (Negative Rumination), and rumination in response to positive mood and progress (Positive Rumination). The second part assesses the products and consequences of rumination, or whether respondents view their ruminative thinking as helpful or not, and consists of three subscales: Constructive Resolution, Ongoing Unresolution, and Move On/Put Behind Me. The third part examines the processes of ruminative thinking and includes four subscales: Comparative Negative Rumination or Affective Interlock, Expansive/Dendritic Thinking, Analytic-Evaluative Abstract Thinking, and Rapid Non-Analytic/Experiential Thinking. Because MBCT theory highlights repetitive negative thinking as a key mechanism underlying the intervention's effects (Segal et al., 2002, 2013), we only used the CERTS Negative Rumination subscale. This subscale consists of 20 items measuring the frequency, duration, controllability, and repetition of rumination in response to five negative contexts (when I feel sad/angry/anxious, when I am by myself, and when I experience a setback on something I value). Items are rated on a 4-point Likert scale ranging from 1 (almost never) to 4 (almost always), with subscale scores ranging from 20 to 80. Total scores provide a general index of the severity of rumination, with higher scores indicating greater negative rumination. In the current sample, only participants from the PREVENT trial (n = 154) completed this measure. Cronbach's alphas for the Negative Rumination subscale in this sample at baseline and post-MBCT were .82 and .83, respectively.

Procedure

Participants completed the FFMQ as well as other measures both before and after MBCT. Measures were administered by research assistants blind to group allocation. The MBCT programme integrates intensive mindfulness meditation practice with psychological education from cognitive behavioural therapy for depression (Segal et al., 2002, 2013). The programme in both PREVENT and SWAD trials consisted of eight weekly 2- to 2.25-hour group sessions and followed the manualised MBCT intervention described by Segal et al. The groups were delivered by therapists who had met the Mindfulness-Based Interventions Teacher Assessment Criteria (Crane et al., 2013), to ensure that the sessions were delivered to a high standard and adhered to the MBCT manual. Written informed consent was obtained from all participants.

Table 8.

Descriptive statistics, reliability coefficients, and sensitivity to change statistics for FFMQ-39 and FFMQ-15 facets pre- and post-MBCT (N = 238)

Scale and subscale	Pre-MBCT			Post-MBCT			Sensitivity to change	
	M	*SD*	α	*M*	*SD*	α	*t* [a]	*d* [95% CI][b]
FFMQ-39								
Observing	25.00	5.78	.78	28.32	5.02	.82	-10.14*	-0.61 [-0.74, -0.48]
Describing	26.21	6.36	.88	27.79	6.11	.90	-5.28*	-0.25 [-0.35, -0.16]
Acting with awareness	24.12	5.29	.84	25.87	4.93	.86	-5.40*	-0.34 [-0.47, -0.21]
Non-judging of experience	24.72	6.12	.86	27.71	5.85	.88	-7.56*	-0.50 [-0.63, -0.36]
Non-reactivity to experience	20.10	4.94	.83	22.70	4.28	.85	-7.80*	-0.56 [-0.71, -0.41]
FFMQ-15								
Observing	8.98	2.73	.64	10.37	2.26	.69	-8.78*	-0.55 [-0.68, -0.42]
Describing	9.84	2.74	.80	10.55	2.60	.83	-5.20*	-0.26 [-0.36, -0.16]
Acting with awareness	9.10	2.25	.68	10.11	2.09	.70	-6.98*	-0.47 [-0.61, -0.33]
Non-judging of experience	9.43	2.67	.76	10.71	2.44	.78	-7.31*	-0.50 [-0.64, -0.36]

Non-reactivity to experience	8.58	2.30	.66	9.68	2.10	.77	-6.47*	-0.50 [-0.65, -0.34]

Note. CI = confidence interval; FFMQ = Five-Facet Mindfulness Questionnaire; MBCT = Mindfulness-Based Cognitive Therapy. For FFMQ-39 subscales, scores range from 8 to 40, except for non-reactivity to experience scores, which range from 7 to 35. For all FFMQ-15 subscales, scores range from 3 to 15. *$p < .001$.

[a] Paired-samples *t*-tests comparing pre-MBCT and post-MBCT scores on the FFMQ facets.

[b] Cohen's *d* effect sizes, with accompanying 95% CIs.

Statistical Analyses

Preliminary analyses. Preliminary analyses were conducted to check for univariate and multivariate normality, and to report the descriptive statistics and general psychometric properties of the FFMQ-39 and FFMQ-15. Cronbach's alpha reliability coefficients were computed for subscales from both versions of the FFMQ. Sensitivity to change pre- to post-MBCT was also examined for FFMQ-39 and FFMQ-15 facets using paired-samples *t*-tests and accompanying Cohen's *d* effect sizes and 95% confidence intervals (CIs) for *d*, calculated using equations 4, 15 and 18 from Nakagawa and Cuthill (2007):

(4) $d = t_{\text{paired}} \sqrt{\frac{2(1 - r_{12})}{n}}$

(15) 95% CI = $ES - 1.96se$ to $ES + 1.96se$

(18) $se_d = \sqrt{\frac{2(1- r_{12})}{n} + \frac{d^2}{2(n-1)}}$

where *d* and *ES* = Cohen's *d* effect size; t_{paired} = the *t* value from the paired *t*-test; r_{12} = the correlation coefficient between the two groups; $n = n_1 = n_2$; CI = confidence interval; *se* = the asymptotic standard error for *d*.

Pearson correlation coefficients (*r*) were conducted to examine the relationships between the FFMQ-39 and FFMQ-15 facets at baseline. As data for both long and short versions of the FFMQ were based on a single administration of the measure, we additionally computed Levy's (1967) corrected correlation coefficients (r_c) in order to adjust for overlapping error variance and spuriously inflated correlations between the long and short versions. Levy's corrected correlations were calculated using the Shortform Version 1.1 software developed by Barrett (2005).

Pearson correlations were also calculated between FFMQ-39 and FFMQ-15 total facet scores and total scores on the BDI-II and the negative rumination subscale of the CERTS before and after MBCT to examine convergent validity. In order to determine whether correlation coefficients with the BDI-II and CERTS were different in size for both versions of the FFMQ, Steiger's (1980) *Z*-tests were conducted. This is the recommended test for comparing two correlations with one variable in common from the same sample (Meng, Rosenthal, & Rubin, 1992). These tests determined whether there were statistically significant differences between FFMQ-39 and FFMQ-

15 facets in their correlations with the BDI-II and CERTS before and after MBCT (e.g., whether there was a significant difference between the correlation coefficient for FFMQ-39 describing and BDI-II pre-MBCT and the correlation for FFMQ-15 describing and BDI-II pre-MBCT). Steiger's *Z*-tests were conducted using software developed by Lee and Preacher (2013). As this test was conducted 20 times (comparing each of the five facet's correlation with BDI-II and CERTS at baseline and post-MBCT across the two versions), Bonferroni-corrected alpha levels of $p = .0025$ (.05/20) were used. Excluding those conducted using the Shortform 1.1 and Steiger's *Z* software, all preliminary analyses were carried out with SPSS Version 22 (IBM, 2013).

Confirmatory factor analyses. To replicate Baer et al. (2006) and M. J. Williams et al. (2014), the following five models were tested separately for both the short and long versions of the FFMQ before and after MBCT using CFA: (1) a one-factor model in which all items are indicators of an overall, latent mindfulness factor, (2) a five-factor model in which items are indicators of their respective five correlated mindfulness factors, (3) a five-factor hierarchical model in which the five factors are indicators of an overarching mindfulness factor, (4) a four-factor hierarchical model in which four factors (minus the observing facet) are indicators of an overarching mindfulness factor, and (5) a four-factor model in which items are indicators of their respective four correlated mindfulness factors (minus the observing facet). To replicate the procedure used by Baer et al. (2006; Baer et al., 2008) and M. J. Williams et al., the CFAs of the 39-item FFMQ were conducted using item parcels (groups of items) rather than individual items.

Following Baer et al. (2008), within each facet, items were allocated sequentially to parcels in the order that they appear on the FFMQ (i.e., first item to Parcel 1, second to Parcel 2 etc.) and item scores within each parcel were averaged. A total of 15 parcels (three parcels per facet, with two or three items per parcel) were used for the CFAs of the FFMQ-39. Item parceling is a controversial practice with several advantages and disadvantages (see Little, Cunningham, Shahar, & Widamon, 2002, for a discussion of the strengths and limitations of parceling). One strength of parceling is that the reliability and stability of a parcel as an indicator of a latent construct tends to be greater than that of an individual item. However, opponents of parceling have argued that parcels can obscure model misspecifications, by improving model fit irrespective of whether or not the model is correctly specified. Considering both the pros and cons of parceling, Little et al. concluded that this technique can be particularly

effective when items within parcels are unidimensional, or measuring the same construct. Other researchers have also stated that parceling should be considered only when there is unidimensionality (e.g., Bandalos & Finney, 2001). In the case of the FFMQ, unidimensionality of its items has been previously established using EFA (Baer et al., 2006). Thus, parceling was deemed appropriate for the current study. Because the 15-item FFMQ consists of just three items per facet, which would be unfeasible to parcel, the individual items were used in the CFAs for this version. In all separate CFA models, error terms were not allowed to covary and items or parcels were constrained to load onto only one factor.

The following six fit indices were used collectively to indicate the global fit of the models to the data: the comparative fit index (CFI; Bentler, 1990), the root mean square error of approximation (RMSEA; Steiger, 1990), the non-normed fit index (NNFI; Bentler & Bonett, 1980), the standardised root mean square residual (SRMR), the chi-square model test, and the Akaike information criterion (AIC; Akaike, 1974). Rules of thumb for the cut-off values on the fit indices which indicate acceptable fit are as follows: the CFI and NNFI should be .95 or higher (Hu & Bentler, 1999), the RMSEA should be less than .10 (Browne & Cudeck, 1993), and the SRMR should be less than .10 (Hu & Bentler, 1999). The AIC was used as a measure of model parsimony to compare the fit of the five models; the lower the value, the better the fit. The chi-square test is generally not recommended to evaluate model fit because of its sensitivity to non-normality, large correlations between variables, large sample sizes, and variables with high proportions of unique variance (Kline, 2011). Therefore, we reported the chi-square test alongside alternative fit indices but did not use it as a primary measure of model fit.

Schermelleh-Engel, Moosbrugger, and Müller (2003) cautioned that cut-off criteria can be arbitrary such that a model may provide a good fit to the data even when one or more fit indices suggest poor fit, and vice versa. Therefore, based on the systematic procedure used by M. J. Williams et al. (2014), the following criteria were also considered when interpreting which model provided a superior fit to the data: (1) significant loadings of items, parcels, or facets onto relevant latent factors, (2) significant covariances between facets, and (3) lowest AIC. All CFAs were conducted in SPSS Amos, Version 22.

Factorial invariance. In addition to the separate CFA models conducted for both versions of the FFMQ before and after MBCT, we evaluated longitudinal factorial

invariance, or whether the same construct is assessed across time, using the analytic approach by Widaman, Ferrer, and Conger (2010). Widaman et al. described four levels involved in establishing factorial invariance. The first step involves testing for configural invariance of the five-factor hierarchical models before and after MBCT simultaneously in a single model. Two longitudinal configural invariance models were tested; one for the FFMQ-39 and one for the FFMQ-15. These tests aimed to establish whether the structural configuration (number of factors and pattern of factor loadings) of the FFMQ was equal across the time points. Good global model-data fit would indicate that the structural configuration of the FFMQ remains stable before and after MBCT. Poor fit of a longitudinal configural model would indicate that the five-factor hierarchical structure of the FFMQ does not apply both before and after MBCT. Covariances were included between the overarching mindfulness factors pre- and post-MBCT, and between the pre-MBCT items/parcels and the corresponding items/parcels post-MBCT. Minimum identification constraints were placed on parameters.

Once configural invariance is established, the next steps would be to test for weak factorial invariance (invariant factor loadings across time), then strong factorial invariance (invariant factor loadings and intercepts across time), and finally strict factorial invariance (invariant factor loadings, intercepts, and factor variances across time) (Widaman et al., 2010). This sequence involves gradually increasing the constraints placed on the model parameters. In order for the FFMQ-39 and FFMQ-15 to be measuring the same construct before and after MBCT, strong or strict factorial invariance must be met. If a preceding level of factorial invariance was not supported, we did not proceed to establish the next level(s) by applying further model constraints.

Results

Preliminary Analyses

All 15 parcels of the FFMQ-39 pre- and post-MBCT were normally distributed, as assessed by checking histograms, box plots, and skewness and kurtosis values. The individual items of the FFMQ-15 pre- and post-MBCT were also normally distributed. No outliers were identified in checking the standardised values for all parcels of the FFMQ-39 and individual items of the FFMQ-15 pre- and post-MBCT. However, Mardia's (1985) test indicated that none of the CFA models met the assumption of

multivariate normality. Under non-normal conditions, the chi-square model test statistic tends to be inflated (so correctly specified models are more likely to be rejected) whereas parameter standard errors tend to be underestimated (so parameters are more likely to be significant) (Chou & Bentler, 1995). Bootstrapping is an approach to managing multivariate non-normality which has been found by empirical studies to perform well relative to other approaches (e.g., Nevitt & Hancock, 2001). Bootstrapping methods in Amos adjust both the *p*-value associated with the chi-square test (Bollen-Stine bootstrap method; Bollen & Stine, 1992) and parameter standard errors (90% bias-corrected CIs). Therefore, all models were analysed twice; first using just maximum-likelihood estimation (MLE; assumes multivariate normality) and second using bootstrapping with 2,000 samples. These two approaches yielded different findings in terms of the significance of some chi-square statistics (the Bollen-Stine chi-square *p*-values for the post-MBCT FFMQ-39 and FFMQ-15 four-factor, four-factor hierarchical, and five-factor models were > .05; using MLE all chi-square tests were significant). However, the chi-square model test is not typically recommended to evaluate the model-data fit (Kline, 2011) and we did not use this as a primary measure of model fit.

Descriptive statistics, Cronbach's alpha reliability coefficients, and sensitivity to change statistics for the 39-item and 15-item FFMQ facets before and after MBCT are given in Table 8. Cronbach's alphas for the FFMQ-39 subscales ranged from .78 to .88 pre-MBCT and .82 to .90 post-MBCT, which correspond closely to the values found in previous research (Baer et al., 2006, 2008; M. J. Williams et al., 2014). Internal consistency values for the FFMQ-15 subscales were generally lower, ranging from .64 to .80 pre-MBCT and .69 to .83 post-MBCT. These alphas are consistent with the range found in previous research using the FFMQ-15 (Baer et al., 2012) and are considered adequate for measures of psychological constructs (Kline, 1999). Differences in internal consistencies between the two forms are unsurprising given that alpha increases with the number of items in a measure (Cortina, 1993). The FFMQ-39 and FFMQ-15 were found to be sensitive to change, as indicated by small/moderate to moderate/large significant increases in subscale scores from pre- to post-MBCT for both versions. Correlations between the total facet scores of the FFMQ-15 and FFMQ-39 at baseline were large and significant, indicating that both versions measured highly similar constructs: $r = .87$ for observing ($r_c = .70$), $r = .94$ for describing ($r_c = .85$), $r =$

.85 for acting with awareness (r_c = .71), r = .90 for non-judging (r_c = .80), and r = .91 for non-reactivity (r_c = .75) (ps < .01).

Convergent Validity

Table 9 presents the correlations between baseline total facet scores on the FFMQ-39 and FFMQ-15 and baseline scores on other constructs, and post-MBCT FFMQ-39 and FFMQ-15 facet scores and post-MBCT scores on other measures. Significant small/moderate to large negative correlations were found between facets of the FFMQ-39 and FFMQ-15 and depression (BDI-II) and negative rumination (CERTS) at both time points. Steiger's *Z*-tests showed that there were no significant differences between the FFMQ-39 and FFMQ-15 facets in their correlations with BDI-II and CERTS at both pre- and post-MBCT (e.g., there was no significant difference between the correlation coefficient for FFMQ-39 acting with awareness and BDI-II at baseline [r = -.38] and the correlation for FFMQ-15 acting with awareness and BDI-II at baseline [r = -.26]). This indicates that the size of the relationships between the FFMQ-15 facets and depression/negative rumination did not differ significantly from the size of relationships found between the FFMQ-39 facets and the same constructs, at both pre- and post-MBCT.

Confirmatory Factor Analyses

Table 10 presents the fit indices for the five CFA models tested for the FFMQ-39 pre-MBCT and post-MBCT, and for the FFMQ-15 pre-MBCT and post-MBCT. For each model, bold indices (CFI, RMSEA, NNFI, and SRMR) indicate that they meet the cut-off criteria for acceptable fit. For both versions of the FFMQ, pre- and post-MBCT, all fit indices indicate that a one-factor model was a poor fit to the data, suggesting that items were not directly subsumed under a unidimensional mindfulness construct. For both versions of the FFMQ, all models fit the data better post-MBCT compared to pre-MBCT. Based on the fit indices, the four-factor and five-factor models best fit the FFMQ-39 pre-MBCT data, the four-factor and five-factor models best fit the FFMQ-15 pre-MBCT data, and the four-factor hierarchical and five-factor models best fit the FFMQ-39 and FFMQ-15 post-MBCT data.

Given the arbitrary nature of the cut-off criteria for fit indices (Schermelleh-Engel et al., 2003), the loadings of items, parcels, or facets onto relevant factors and the

relationships between facets were also taken into account. Across all four data sets, in a five-factor hierarchical model, all five facets loaded significantly onto an overall mindfulness factor and in a four-factor hierarchical model, all four facets (minus the observing facet) loaded significantly onto an overarching mindfulness factor (*p*s < .01) (see Table 11). Taking these significant loadings into account, this indicates that facets of both versions of the FFMQ at both time points, can be considered part of an overall mindfulness factor. All loadings of items and parcels onto relevant facets were also significant (*p*s < .01). However, for the FFMQ-39 pre-MBCT data, the covariance between the observing and non-judging facets was non-significant in a five-factor model ($p = .43$); in a four-factor model, all covariances between facets were significant (*p*s < .01). Additionally, for the FFMQ-15 pre-MBCT data, non-significant covariances were found between observing and acting with awareness ($p = .77$), and observing and non-judging ($p = .71$); in a four-factor model, all covariances were significant (*p*s < .05). All covariances between facets were significant in the post-MBCT models for both versions of the FFMQ (*p*s < .01). This suggests that pre-MBCT, both versions of the FFMQ measure four, not five, related facets of mindfulness (excluding the observing facet). Tables 1 to 4 in Appendix D present item loadings and factor loadings for the four-factor hierarchical pre-MBCT model and five-factor hierarchical post-MBCT model for both versions of the FFMQ.

Table 10 also displays the fit indices for the configural invariance models, which tested the five-factor hierarchical models of the FFMQ-39 or FFMQ-15 before and after MBCT simultaneously. The configural invariance models for both versions of the FFMQ fit poorly to the data; almost all of the indices did not meet the cut-off criteria for acceptable fit. As configural invariance was not supported, we did not apply further model restrictions to test for weak, strong, or strict factorial invariance.

Taken together, the pattern of findings suggests that a four-factor hierarchical model provided the optimal fit for both versions of the FFMQ pre-MBCT, whereas a five-factor hierarchical model was superior for both versions of the FFMQ post-MBCT. Tests of configural invariance support this interpretation, by indicating that the structure of both versions of the FFMQ was not equivalent before and after MBCT. Although the fit indices favour the non-hierarchical models, the arbitrary nature of cut-off criteria for these indices (Schermelleh-Engel et al., 2003) coupled with the consideration of other criteria (e.g., significant loadings of facets onto an overarching

mindfulness factor as shown in Table 11) lend strong support for the hierarchical models.

Table 9.

Pearson correlation coefficients between total facet scores on the FFMQ-39 and FFMQ-15 and other constructs pre- and post-MBCT

Scale/subscale	BDI-II		CERTS: Negative rumination[a]	
FFMQ-39	Pre	Post	Pre	Post
Observing	-.15*	-.18**	-.23**	-.28**
Describing	-.28**	-.29**	-.18*	-.27**
Acting with awareness	-.38**	-.50**	-.37**	-.53**
Non-judging of experience	-.32**	-.44**	-.54**	-.64**
Non-reactivity to experience	-.24**	-.36**	-.39**	-.49**
FFMQ-15				
Observing	-.15*	-.18**	-.18*	-.23**
Describing	-.25**	-.29**	-.22**	-.28**
Acting with awareness	-.26**	-.40**	-.35**	-.52**
Non-judging of experience	-.30**	-.42**	-.50**	-.58**
Non-reactivity to experience	-.26**	-.36**	-.36**	-.43**

Note. BDI-II = Beck Depression Inventory–II; CERTS = Cambridge-Exeter Ruminative Thinking Scale; FFMQ = Five-Facet Mindfulness Questionnaire.

[a] This measure was only completed by participants in the PREVENT trial (n = 154). In the PREVENT sample 46 participants (30%) were male and 108 (70%) female. The mean age was 52.12 years (SD = 11.18) and 152 (98.7%) of the sample were white. Ten participants (7%) had no qualifications, 23 (15%) had some GCSE/O-Level qualifications, 64 (42%) had A-Levels or vocational qualifications, 37 (24%) had a bachelor's degree, 8 (5%) had a master's degree, and 10 (6.5%) had a PhD or professional qualification. Data on education were missing for two participants.

*p < .05 **p < .01.

Table 10.

CFA fit indices for the five models tested pre- and post-MBCT for both long (FFMQ-39) and short (FFMQ-15) versions of the FFMQ and the configural invariance models

Model	CFI	RMSEA [90% CI]	NNFI	SR-MR	χ^2	*df*	AIC
FFMQ-39 Pre-MBCT							
One-factor	.455	.225 [.213, .236]	.364	.168	1167.386	90	1227.386
Four-factor	**.950**	**.086 [.069, .103]**	.931	**.052**	131.792	48	191.792
Four-factor hierarchical[a]	.939	**.092 [.076, .110]**	.920	**.071**	151.358	50	207.358
Five-factor	**.951**	**.071 [.057, .085]**	.936	**.050**	175.948	80	255.948
Five-factor hierarchical[b]	.932	**.082 [.069, .095]**	.916	**.085**	219.807	85	289.807
FFMQ-39 Post-MBCT							
One-factor	.508	.230 [.219, .242]	.426	.142	1222.433	90	1282.433
Four-factor	**.987**	**.047 [.023, .068]**	**.982**	**.037**	72.908	48	132.908
Four-factor hierarchical	**.986**	**.047 [.024, .068]**	**.981**	**.042**	76.453	50	132.453
Five-factor	**.987**	**.040 [.018, .057]**	**.983**	**.037**	110.068	80	190.068
Five-factor hierarchical	**.969**	**.060 [.045, .074]**	**.961**	**.068**	157.012	85	227.012
FFMQ-15 Pre-MBCT							
One-factor	.458	.157 [.145, .169]	.367	.127	613.755	90	673.755
Four-factor	.932	**.070 [.052, .089]**	.906	**.076**	103.992	48	163.992
Four-factor hierarchical	.915	**.077 [.059, .094]**	.888	**.084**	119.699	50	175.699
Five-factor	.925	**.062 [.047, .077]**	.902	**.071**	152.158	80	232.158
Five-factor hierarchical	.895	**.071 [.057, .085]**	.870	**.089**	186.794	85	256.794
FFMQ-15 Post-MBCT							
One-factor	.553	.159 [.147, .171]	.478	.118	627.361	90	687.361
Four-factor	**.956**	**.062 [.042, .081]**	.940	**.058**	91.676	48	151.676

Four-factor hierarchical	**.956**	**.061 [.042, .080]**	.942	**.060**	94.213	50	150.213
Five-factor	**.947**	**.058 [.042, .073]**	.930	**.055**	143.624	80	223.624
Five-factor hierarchical	.926	**.067 [.052, .081]**	.908	**.069**	174.442	85	244.442
FFMQ-39 Configural Invariance[c]	.886	**.079 [.073, .085]**	.869	.104	941.057	379	1113.057
FFMQ-15 Configural Invariance	.825	**.073 [.066, .079]**	.800	.096	852.401	379	1024.401

Note. AIC = Akaike information criterion; CFA = confirmatory factor analysis; CFI = comparative fit index; CI = confidence interval; FFMQ = Five-Facet Mindfulness Questionnaire; MBCT = Mindfulness-Based Cognitive Therapy; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual. Indices given in bold indicate that they are within the acceptable range when rounded up or down to two decimal places.

[a] Four-factor hierarchical refers to the model in which the facets describe, acting with awareness, non-judging, and non-reactivity (without the observing facet) loaded onto an overall mindfulness factor.

[b] Five-factor hierarchical refers to the model in which all five facets loaded onto an overall mindfulness factor.

[c] Configural invariance refers to the model in which pre- and post-MBCT five-factor hierarchical models are tested simultaneously. One model was tested for the FFMQ-39 and one for the FFMQ-15.

*$p < .001$.

Table 11.

Standardised loadings of the facets onto an overarching mindfulness factor for the FFMQ-39 and FFMQ-15 four-factor hierarchical and five-factor hierarchical models pre-MBCT and post-MBCT

Path	Pre-MBCT	Post-MBCT
FFMQ-39 four-factor hierarchical		
Describing <- Mindfulness	.63	.57
Acting with awareness <- Mindfulness	.72	.75
Non-judging <- Mindfulness	.53	.74
Non-reactivity <- Mindfulness	.63	.73
FFMQ-39 five-factor hierarchical		
Observing <- Mindfulness	.58	.63
Describing <- Mindfulness	.71	.59
Acting with awareness <- Mindfulness	.56	.65
Non-judging <- Mindfulness	.39	.66
Non-reactivity <- Mindfulness	.75	.86
FFMQ-15 four-factor hierarchical		
Describing <- Mindfulness	.41	.47
Acting with awareness <- Mindfulness	.59	.69
Non-judging <- Mindfulness	.73	.86
Non-reactivity <- Mindfulness	.58	.60
FFMQ-15 five-factor hierarchical		
Observing <- Mindfulness	.35	.55
Describing <- Mindfulness	.54	.51
Acting with awareness <- Mindfulness	.43	.65
Non-judging <- Mindfulness	.53	.72
Non-reactivity <- Mindfulness	.81	.73

Note. FFMQ = Five-Facet Mindfulness Questionnaire; MBCT = Mindfulness-Based Cognitive Therapy. All loadings of facets onto an overarching mindfulness factor were significant (*p*s < .01).

Discussion

The primary aim of this study was to examine the factor structure of the long (39 item) and short (15 item) versions of the FFMQ before and after MBCT, to determine whether the structure remains stable over a period in which people are learning mindfulness meditation. The secondary aim was to assess the general psychometric properties of the FFMQ-39 and FFMQ-15, specifically their sensitivity to change and convergent validity before and after MBCT. We found both versions to be sensitive to change; small/moderate to moderate/large significant increases from pre- to post-MBCT were found for total facet scores from both versions. Additionally, large correlations were found between the total facet scores of the FFMQ-15 and FFMQ-39 (r ranged from .85 to .94, r_c ranged from .70 to .85), indicating that both versions measured highly similar constructs. Convergent validity was tested by correlating FFMQ total facet scores with theoretically related constructs (depression and negative rumination) before and after MBCT. Significant negative correlations were found between rumination/depression and facets of both versions of the FFMQ. Differences in the correlation coefficients between the two versions of the FFMQ were also found to be non-significant; the size of the relationships between the FFMQ-15 facets and depression/rumination did not differ significantly from the size of relationships found between the FFMQ-39 facets and these variables.

Separate CFAs showed that a four-factor hierarchical model, without the observing facet, provided the best fit for both versions of the FFMQ pre-MBCT, whereas a five-factor hierarchical model was superior for both versions of the FFMQ post-MBCT. This was informed by non-significant covariances between observing and other facets (non-judging, acting with awareness) in the FFMQ-39 and FFMQ-15 pre-MBCT models which were significant in the post-MBCT models. Significant loadings of the facets to a hierarchical latent mindfulness construct also contributed to this interpretation. Additionally, configural invariance was not supported for both versions of the FFMQ; a single model of the five-factor hierarchical structure before and after MBCT corresponded poorly to the data. Taken together, this indicates that the FFMQ's structural configuration, or the number of factors and pattern of factor loadings, was not equivalent across the two time points.

Our CFA findings support the emerging body of literature that has shown that the five-factor hierarchical structure of the FFMQ holds in samples of meditators or

people who have undertaken an MBI, but a four-factor hierarchical model best represents data from people with little or no meditation experience (Baer et al., 2006, 2008; Curtiss & Klemanski, 2014a; M. J. Williams et al., 2014). Additionally, the non-significant covariances found between observing and non-judging (for the FFMQ-39 and FFMQ-15) and observing and acting with awareness (for the FFMQ-15) at baseline reflect the non-significant relationships between observing and other facets found in previous studies (Baer et al., 2006; Bohlmeijer et al., 2011; Curtiss & Klemanski, 2014a). Importantly, our results also extend previous research, by demonstrating that the factor structure of the FFMQ varies before and after MBCT as well as across samples of meditators and non-meditators. Furthermore, our findings highlight that the FFMQ-39 and FFMQ-15 are both sensitive to change, are consistent in terms of factor structure before and after MBCT, and do not differ significantly from each other with regard to convergent validity.

In relation to the observing subscale, current findings support the perspective that meditation experience alters people's qualities of noticing, by enhancing the strength of the relationship between observing experience and other aspects of mindfulness (in particular acting with awareness and non-judging) (Baer et al., 2008). It is possible that participants with little or no meditation experience report how much they tend to observe, but the way in which they observe may not be consistent with mindfulness and may instead involve neutral or maladaptive forms of attention. With meditation experience and familiarity with a more accepting and curious way of noticing all experience, not only may people report greater levels of observing, but the way in which they observe may be more consistent with acting with awareness and non-judging. For example, observing a negative thought such as "this happy moment won't last" may be associated with noting and letting the thought pass, while reorienting attention back to the present moment and to other dimensions of experience.

While current findings indicate that people's quality of observing differs before and after MBCT, they do not provide direct support for the explanation that observing may involve pathological forms of attentional monitoring pre-MBCT and accepting, curious, and purposeful attention post-MBCT. To test this, studies would need to examine whether the relationship between pre-MBCT scores on the observing subscale and anxious monitoring is significantly greater than the correlation between post-MBCT observing scores and anxious monitoring. Alternatively, studies could compare

pre-MBCT observing scores between non-meditators with a diagnosis of an anxiety disorder, who have a higher degree of anxious monitoring, and non-meditators in the non-clinical population or without a diagnosis of anxiety. Once a better understanding is reached regarding how the observing subscale operates, it may be useful to explore the effects of particular facets (e.g., acting with awareness) on the functioning of the observing facet (e.g., Desrosiers, Vine, Curtiss, & Klemanski, 2014).

Implications

Several implications arise from our findings for studies investigating change in trait mindfulness, as measured by the FFMQ, before and after mindfulness interventions. Our findings show that total FFMQ scores and scores on the observing facet are not valid for evaluating change from pre to post intervention; pre to post differences in scores on the observing subscale may reflect changes in the extent to which people notice experience, rather than a genuine change in the ability to observe mindfully. To evaluate interventions that involve mindfulness meditation practice, researchers should consider only comparing the describing, acting with awareness, non-judging, and non-reactivity facets and combine only these four subscale scores into a total FFMQ score. Although the empirical evidence suggests that for non-meditators the observing facet does not converge well with other facets which underlie mindfulness, this does not mean that theoretically, observing experience is not an integral aspect of a mindful disposition. Rather, this suggests that although the current observing items may reflect how much people tend to notice, they may need revision to better capture the accepting, curious, and purposeful qualities of noticing all experience consistent with a mindful disposition. Future research should also consider using a triangulated approach, whereby alternative methods of measuring mindfulness (e.g., neuropsychological, cognitive, and qualitative measures; see Sauer et al., 2013, for a review) are used to complement the FFMQ.

Additionally, current findings support the FFMQ-15 as a valid and reliable alternative measure to the original FFMQ for use in studies administering multiple measures and/or questionnaires at multiple time points. Furthermore, the significant loadings of all FFMQ-39 and FFMQ-15 facets onto an overall mindfulness factor pre- and post-MBCT supports the legitimacy of using a total FFMQ score, alongside total facet scores, as an indicator of global mindfulness level (but without the observing

facet in non-meditating samples). Support for the hierarchical models of the FFMQ also reinforces the theoretical conceptualisation of mindfulness as a multifaceted yet coherent construct.

Limitations and Future Directions

Current findings inform our use of the FFMQ to measure mindfulness, which is essential to advancing research in this area. However, there are several limitations. Prior mediation experience was not measured in the current sample. Participants may have had experience of mindfulness meditation prior to the MBCT programme, which would question the validity of attributing changes in the factor structure of the FFMQ to learning mindfulness. It is possible that changes in the factor structure might occur across other types of psychological intervention, following changes in level of depression within the sample, or as a result of the passage of time and re-testing. However, our findings showing that a four-factor hierarchical model and a five-factor hierarchical model best fit the data pre-MBCT and post-MBCT, respectively, support previous research conducted in non-meditator and mediator samples (Baer et al., 2006, 2008; Curtiss & Klemanski, 2014a; M. J. Williams et al., 2014). We also found non-significant covariances between observing and non-judging and acting with awareness facets, which correspond closely with previous findings (Baer et al., 2006; Bohlmeijer et al., 2011; Curtiss & Klemanski, 2014a). These parallels suggest that our sample is likely to have had a similar level of meditation experience pre-MBCT as non-meditator samples. Nonetheless, meditation status should be recorded in future studies as this will allow replication of previous research, by conducting multiple group CFAs to assess whether baseline FFMQ factor structure is altered by meditation experience.

The current sample was also limited to adults with MDD in remission, the population for whom MBCT was originally developed. The present findings should be extended by testing the FFMQ models on pre- and post-MBCT data from other clinical and non-clinical (e.g., students, community samples) samples. We would expect our findings to be replicated in independent samples, provided that participants have little or no meditation experience before MBCT. The factor structure and psychometric properties of the FFMQ-15 should also be tested in additional samples to further support its use. In line with methodological guidelines for the development and validation of short form measures (Smith, McCarthy, & Anderson, 2000), we

recommend that studies conducting additional psychometric testing of the FFMQ-15 administer the short form in its own right to an independent sample (i.e., not a sample in which the FFMQ-39 was administered). In the current study, data for both long and short versions of the FFMQ were based on a single administration of the measure and correlation coefficients were corrected to account for overlapping error variance. However, we recommend that future research examining the overlap of the FFMQ-15 and FFMQ-39 administer both versions, independently, to the same participants.

A further limitation of this study pertains to its sample size ($N = 238$). In the most complex separate CFA model (five-factor correlated model) there were 40 free parameters and in the configural invariance models there were 86 free parameters. The common rule of thumb of at least five participants per free parameter (Bentler & Chou, 1987) would mean that the current sample size may have been adequate for separate CFA analyses but not for analyses of configural invariance. However, it is widely acknowledged that rules of thumb for determining sample size requirements do not apply to all situations and need to take into account additional factors such as degrees of factor overdetermination and item communalities (Meade & Bauer, 2007). Nevertheless, it would be desirable for future studies to replicate the configural invariance analyses using larger clinical and non-clinical samples.

Moreover, future research could assess whether current findings are replicated using other interventions which involve substantial mindfulness meditation practice, such as MBSR, and interventions which include mindfulness principles but less or no meditation practice, such as acceptance and commitment therapy (Hayes & Wilson, 1994) and briefer self-help MBIs (see Cavanagh, Strauss, Forder, & Jones, 2014 for a review). This could potentially yield interesting insights into the degree to which meditation practice is needed to alter the way we observe experience and whether changes in the factor structure of the FFMQ are caused specifically by meditation practice, or by other factors (e.g., intellectual understanding of mindfulness).

Conclusion

The FFMQ is a widely used measure of dispositional mindfulness in studies investigating change before and after MBIs, such as MBCT and MBSR. However, our findings show that the factor structure of the FFMQ is not invariant before and after MBCT and suggest that researchers should consider omitting the observing subscale

when comparing total scale and subscale scores before and after mindfulness interventions. Current findings also provide initial support for the 15-item version of the FFMQ as a reliable and valid alternative measure for use in studies administering multiple measures and/or measures at multiple occasions.

Chapter 5:
What is Compassion and How Can We Measure it? A Review of Definitions and Measures

Chapter 5 is published in *Clinical Psychology Review* as:

Strauss, C., Taylor, B. L., **Gu, J.**, Kuyken, W., Baer, R., Jones, F., & Cavanagh, K. (2016). What is Compassion and How Can We Measure it? A Review of Definitions and Measures. *Clinical Psychology Review, 47*, 15-27.

Author Contributions

Design:	All authors
Data Collection:	JG and BLT
Data Analysis:	JG and BLT
Initial Draft:	JG, BLT, and CS
Reviewing and Editing:	All authors

Abstract

The importance of compassion is widely recognised and it is receiving increasing research attention. Yet, there is lack of consensus on definition and a paucity of psychometrically robust measures of this construct. Without an agreed definition and adequate measures, we cannot study compassion, measure compassion, or evaluate whether interventions designed to enhance compassion are effective. In response, this paper proposes a definition of compassion and offers a systematic review of self- and observer-rated measures. Following consolidation of existing definitions, we propose that compassion consists of five elements: recognising suffering, understanding the universality of human suffering, feeling for the person suffering, tolerating uncomfortable feelings, and motivation to act/acting to alleviate suffering. Three databases were searched (Web of Science, PsycInfo, and Medline) and nine measures included and rated for quality. Quality ratings ranged from 2 to 7 out of 14 with low ratings due to poor internal consistency for subscales, insufficient evidence for factor structure and/or failure to examine floor/ceiling effects, test-retest reliability, and discriminant validity. We call for empirical testing of our five-element definition, and if supported, the development of a measure of compassion based on this operational definition which demonstrates adequate psychometric properties.

Keywords: compassion, self-compassion, measure, systematic review, definition

Introduction

The importance of compassion is recognised in many segments of society. Most of the world's religious traditions place compassion at the centre of their belief systems. International professional bodies in healthcare, education, and the justice system also emphasise the importance of compassion. In the US, compassion is enshrined in the American Medical Association's (AMA) Principles of Medical Ethics, with Item 1 stating that "a physician shall be dedicated to providing competent medical care, with compassion and respect for human dignity and rights" (AMA, 2001). In the UK, compassion is one of the six core values in the NHS constitution (Department of Health; DoH, 2013), and calls for a greater focus on compassion have been driven in part by high profile exposés of serious failings in compassionate care at some hospitals and care homes. The international 'Compassion in Education' foundation (CoED, 2014) offers a range of services to educational professionals in order to promote compassion in the education system. It has also been argued that compassion should lie at the core of the ethical framework guiding our justice systems (Norko, 2005).

An evolutionary perspective on compassion can be traced to Darwin (1871), who stated that "those communities which included the greatest number of the most sympathetic members would flourish best and rear the greatest number of offspring" (p. 130). Current theorists also note that compassion is reproductively advantageous, being part of the care-giving system that has evolved to nurture and protect the young (e.g., Gilbert, 2005; Goetz, Keltner, & Simon-Thomas, 2010). Compassion can be seen as having evolved from an adaptive focus on protecting oneself and one's offspring to a broader focus on protecting others including and beyond one's immediate kinship group (de Waal, 2009). Compassion may also have evolved in primates because it is a desirable criterion in mate selection and facilitates cooperative relationships with non-kin (e.g., de Waal, 2009; Keltner, 2009).

Within the healthcare domain, compassion is believed to have numerous practical advantages. It has been argued that treating patients compassionately has wide-ranging benefits, including improving clinical outcomes, increasing patient satisfaction with services, and enhancing the quality of information gathered from patients (Epstein et al., 2005; Redelmeier et al., 1995; Sanghavi, 2006). Conversely, compassion fatigue may contribute to poor quality of care (Najjar, Davis, Beck-Coon, & Carney Doebbeling, 2009). Treating oneself and others with compassion is also

believed to promote individual wellbeing and improve mental health (e.g., Cosley et al., 2010; Feldman & Kuyken, 2011; MacBeth & Gumley, 2012). Accordingly, some researchers have called for the implementation of interventions that seek to enhance people's ability to give and receive compassion (e.g., Gilbert, 2005, 2010), arguing that compassion buffers reactivity to stress and is central to the process of recovery from psychopathology. Other research has focused on the developmental trajectory of compassion and has found relationships between parenting styles and children's levels of sympathy and caring (Eisenberg, VanSchyndel, & Hofer, 2015) and between attachment security in childhood and capacity for compassion in adulthood (see Gillath, Shaver, & Mukilincer, 2005, for a review).

Despite the importance of compassion and increasing interest from researchers, clinicians, teachers, and other professionals, there is lack of consensus on its definition and a paucity of psychometrically robust measurement tools. Without these, scientific enquiry is greatly impeded – we need consensus on a definition and valid and reliable measurement tools in order to assess compassion in empirical research. This paper has two aims: first, to suggest a definition of compassion based on a consolidation of conceptualisations and definitions in the field and second, to systematically review self- and observer-rated measures of compassion.

Conceptualisations of Compassion: Towards a Definition

According to the Oxford English Dictionary, the word 'compassion' stems from the Latin *compati*, meaning "to suffer with". In the literature, there appears to be a broad consensus that compassion involves feeling for a person who is suffering and being motivated to act to help them (e.g., Lazarus, 1991; Goetz et al., 2010). For example, in his seminal work on human emotions Lazarus defines compassion as "being moved by another's suffering and wanting to help" (p. 289). Similarly, in a major systematic review of compassion and its evolutionary origins, Goetz et al. define it as "the feeling that arises in witnessing another's suffering and that motivates a subsequent desire to help" (p. 351). These definitions have in common the suggestion that compassion is not only about feeling touched by a person's suffering, but also about wanting to act to help them. Compassion is a fundamental tenet of Buddhist philosophy (it is, in fact, emphasised by all the main world religions but Buddhist perspectives on compassion have been given greater prominence in the psychological

literature) and the Dalai Lama (1995) defines compassion in comparable terms as "an openness to the suffering of others with a commitment to relieve it". However, within Buddhism, compassion is seen not only as an emotional response but also as a response founded on reason and wisdom which is embedded in an ethical framework concerned with the selfless intention of freeing others from suffering.

More specifically, in their review of compassion within organisations, Kanov et al. (2004) argue that compassion consists of three facets: noticing, feeling, and responding. 'Noticing' involves being aware of a person's suffering, either by cognitively recognising this suffering or by experiencing an unconscious physical or affective reaction to it. 'Feeling' is defined as responding emotionally to that suffering and experiencing 'empathic concern' through adopting the person's perspective and imagining or feeling their condition. Finally, 'responding' involves having a desire to act to alleviate the person's suffering. As in Buddhist conceptualisations, this definition suggests that compassion does not purely consist of affective and behavioural elements, but also may have cognitive components insofar as it involves being able to imagine and reason about a person's experiences.

Gilbert (2010) conceptualises compassion in evolutionary terms, arguing that compassion is an evolved motivational system designed to regulate negative affect, where compassion is seen to have originated from the same capacities that primates evolved to form attachment bonds and engage in affiliative and cooperative behaviours for group survival. He defines compassion as "a deep awareness of the suffering of another coupled with the wish to relieve it" (Gilbert, 2009a, p. 13) and, like Kanov et al. (2004), suggests it has cognitive, affective and behavioural elements. Gilbert (2010) sees compassion as consisting of six 'attributes': sensitivity, sympathy, empathy, motivation/caring, distress tolerance, and non-judgement. 'Sensitivity' involves being responsive to other people's emotions and perceiving when they need help, which appears to correspond to Kanov et al.'s 'noticing' facet. 'Sympathy' (defined as showing concern for the other person's suffering) and 'empathy' (defined as putting yourself in their shoes) together appear to correspond to Kanov et al.'s 'feeling' facet. Finally, 'motivation' to act is akin to Kanov et al.'s 'responding' facet.

The final two components in Gilbert's (2010) model – 'distress tolerance' and 'non-judgement' – are not included in Kanov et al.'s (2004) model. Distress tolerance is defined as the ability to tolerate difficult emotions in oneself when confronted with someone else's suffering without becoming overwhelmed by them. Gilbert argues that

this is important because if we over-identify with a person's suffering we may feel a need to get away from them or reduce our awareness of their distress, preventing a compassionate response. This suggests that, although compassion is about 'suffering with' another person, if we feel such extreme personal distress in the face of another's suffering that we become too focused on our own discomfort, this may hinder our ability to help. The final element of Gilbert's model – 'non-judgement' – is defined as the ability to remain accepting of and tolerant towards another person even when their condition, or response to it, gives rise to difficult feelings in oneself, such as frustration, anger, fear, or disgust. The idea that compassion means approaching those who are suffering with non-judgement and tolerance – even if they are in some sense disagreeable to us – is also central to Buddhist conceptualisations. For example, the Dalai Lama (2002b) contends that: "for a practitioner of love and compassion, an enemy is one of the most important teachers. Without an enemy you cannot practice tolerance, and without tolerance you cannot build a sound basis of compassion" (p. 75).

Both Gilbert (2005, 2010) and the Dalai Lama are also clear that compassion is not only felt for close others (where attachment comes into play as well), but also for those we do not know. Similarly, Gilbert (2003, cited in Wang, 2005) notes that "one can feel compassion for those we might never meet (the starving children in Africa)" (p. 99-100). The idea that compassion can be experienced towards close others and those we do not know is also emphasised by Sprecher and Fehr (2005) who developed a measure of 'compassionate love' which includes separate versions relating to close others and strangers or humankind at large.

Like Gilbert (2010), Wispe (1991) conceptualises compassion for others not only as being aware of and moved by suffering and wanting to help, but also as including the ability to adopt a non-judgmental stance towards others and to tolerate one's own distress when faced with other people's suffering. Neff (2003a) developed this definition of compassion for others into a model of compassion for the self, arguing that self-compassion can be viewed as compassion directed inwardly towards the self. She concludes that self-compassion consists of three principal components: kindness (being kind and non-judgmental towards the self rather than self-critical), mindfulness (which like 'distress tolerance', involves holding painful feelings in mindful awareness rather than over-identifying with them), and common humanity (seeing one's suffering as part of the human condition rather than as isolating).

It is debatable whether compassion for others and self-compassion are in fact part of the same overarching construct. Whilst Buddhist thinking argues that differentiating compassion for others from self-compassion means drawing a false distinction between the self and others, and moreover that self-compassion is a prerequisite for showing 'true' compassion towards others, recent research has found that associations between self-compassion and other-focused compassion may be weak, or even non-existent for some populations. For example, Neff and Pommier (2013) explored the relationship between self-compassion and compassion for others and found that they were not correlated in a sample of undergraduates (r = .00), and only weakly correlated in a community sample and a sample of practicing meditators (r = .15 and .28, respectively). Similarly, Pommier (2010) found no association between self-compassion and compassion for others in a sample of undergraduates (r = .07). It is unclear whether the lack of association between self-compassion and compassion for others reflects a genuine independence between these two constructs or whether it reflects definitional problems, weaknesses of correlational study designs, or limitations with current measures (e.g., M. J. Williams, Dalgleish, Karl, & Kuyken, 2014). This is an area for further empirical research.

While acknowledging some of the difficulties with equating self-compassion with compassion for others, Pommier (2010) applied Neff's (2003a) model of self-compassion to a model of compassion for others suggesting that, like self-compassion, compassion for others can be seen as involving kindness, mindfulness and common humanity. In Pommier's model, 'kindness' is defined as being understanding towards others who are suffering instead of being critical or indifferent towards them. 'Mindfulness' is seen as the ability to notice another person's suffering and remain open to it without feeling so distressed that you disengage from that person. And 'common humanity' is conceptualised as realising that all humans suffer and that one could find oneself in the position of the sufferer if one was less fortunate – a sense that "There but for the grace of God, go I".

This emphasis on seeing a 'common humanity' with the person who is suffering is also evident in Buddhist definitions of compassion, with the Dalai Lama (2005) arguing that: "Genuine compassion must have both wisdom and loving kindness. That is to say, one must understand the nature of the suffering from which we wish to free others (this is wisdom), and one must experience deep intimacy and empathy with other sentient beings (this is loving kindness)" (p. 49). Within such Buddhist

conceptualisations, understanding the nature of suffering ('wisdom') is to understand that suffering is part of what it is to be human; suffering is a shared human experience. Similarly, in their review of the role of compassion in mindfulness-based therapies, Feldman and Kuyken (2011) describe compassion as "an orientation of mind that recognises pain and the universality of pain in human experience and the capacity to meet that pain with kindness, empathy, equanimity, and patience" (p. 145).

In summary, in all these definitions compassion is seen as awareness of someone's suffering, being moved by it (emotionally and, according to some definitions, cognitively), and acting or feeling motivated to help. Several definitions emphasise that, although one is moved by suffering, compassion also involves being able to tolerate uncomfortable feelings that arise in oneself as a result of seeing suffering, including tolerating feelings of distaste, frustration, or anger that might be elicited by that suffering. There is also a suggestion in several definitions that compassion involves recognising a commonality with the sufferer, acknowledging that as a fellow being we too could find ourselves in a similar position. Table 12 contains summaries of the major definitions of compassion discussed in this section.

Related Constructs

In definitions of compassion, reference is commonly made to related terms such as empathy and in turn these words are often used to define each other. The similarities between compassion and constructs such as kindness, pity, and altruism have also been noted (Goetz et al., 2010). It is instructive to consider the overlaps and distinctions between these terms.

According to the Oxford English Dictionary, the word empathy is defined as "the power of mentally identifying oneself with (and so fully comprehending) a person or object of contemplation". Like compassion, empathy has been described as a multidimensional construct, consisting of cognitive and affective components (Davis, 1983). Cognitive empathy can be defined as intellectually understanding another person's emotions and perspective (Hogan, 1969), whereas affective empathy refers to being affected by and sharing another's emotions (Mehrabian & Epstein, 1972). Gilbert (2010), Kanov et al. (2004), and the Dalai Lama (2005) all explicitly define compassion as requiring empathy and therefore appear to see empathy as an essential element of compassion. Even so, they suggest that compassion has additional

components over and above empathy. In particular, a desire to act or acting to alleviate suffering is seen as a core feature of compassion but not empathy (see Table 12).

A second distinction between compassion and empathy is that whereas compassion is felt specifically in response to suffering, empathy may apply to a broader range of situations, for example one could feel empathy with someone else's anger, fear, or joy (Pommier, 2010). Moreover, Goetz et al. (2010) argue that compassion is an emotion in its own right, whereas empathy is the vicarious experience of another's emotions, while Sprecher and Fehr (2005) contend that compassion is broader than empathy because it can be felt for humanity at large, rather than only in relation to specific interpersonal encounters. In addition, recent neuroscientific findings suggest that different brain regions are activated in response to compassion and empathy training (Klimecki, Leiberg, Ricard, & Singer, 2013).

The same is true of pity, which, despite also having similarities to compassion, does not require an inclination to help. On the contrary, some writers have argued that pity implies that one sees someone as unworthy of help (Lazarus, 1991), or at least involves showing condescension towards them (e.g., Cassell, 2002). At the other end of the spectrum, altruism has a greater focus than compassion on behavioural acts that may be at a great personal cost to the person. Altruistic acts can also have a broad range of motivations that do not necessarily involve the same elements as compassion.

Finally, compassion is frequently linked to kindness (defined by the Oxford English Dictionary as "the quality of being friendly, generous, and considerate"). For example, Neff (2003a) and Pommier (2010) include 'kindness' as a component of compassion and compassion has even been defined as "intelligent kindness" (e.g., DoH, 2013). However, these two terms also have distinctions. For example, as outlined, compassion includes elements beyond kindness (e.g., recognising and being touched by suffering) and likewise, kindness includes elements beyond compassion, as kindness is not only linked to suffering (e.g., remembering someone's birthday is kind but not compassionate). Additionally, compassion may not always involve kindness in the moment (e.g., taking a 'tough love' approach may be compassionate but not kind).

Table 12.

Major definitions of compassion in the literature in relation to the five-element definition of compassion

Definition of Compassion	Recognising suffering[a]	Understanding the universality of suffering	Emotional resonance	Tolerating uncomfortable feelings	Motivation to act/acting to alleviate suffering
1. "Being moved by another's suffering and wanting to help" (Lazarus, 1991, p. 289).	✓ (implied)		✓		✓
2. An openness to the suffering of others with a commitment to relieve it (Dalai Lama, 1995). Buddhist conceptualisations also highlight cognitive components (e.g. the ability to imagine and reason about a person's experiences) and approaching those who are suffering with tolerance and non-judgement.	✓ (implied)	✓	✓	✓	✓
3. "Being touched by the suffering of others, opening one's awareness to others' pain and not avoiding or disconnecting from it, so that feelings of kindness towards others and the desire to alleviate	✓ (explicitly stated)		✓	✓	✓

their suffering emerge. It also involves offering non-judgmental understanding to those who fail or do wrong" (Neff, 2003a, p. 86-87).					
4. Compassion consists of three facets: Noticing, feeling, and responding (Kanov et al., 2004).	✓ (explicitly stated)		✓		✓
5. "A deep awareness of the suffering of another coupled with the wish to relieve it" (Gilbert, 2009a, p. 13). Compassion consists of six 'attributes': Sensitivity, Sympathy, Empathy, Motivation/Caring, Distress Tolerance, and Non-Judgement.	✓ (explicitly stated)		✓	✓	✓
6. "The feeling that arises in witnessing another's suffering and that motivates a subsequent desire to help" (Goetz et al., 2010, p. 351).	✓ (explicitly stated)		✓		✓
7. "An orientation of mind that recognises pain and the universality of pain in human experience and the capacity to meet that pain with kindness, empathy, equanimity and patience" (Feldman & Kuyken, 2011, p. 145).	✓ (explicitly stated)	✓	✓	✓	
8. Compassion involves three elements: Kindness, mindfulness, and common humanity (Pommier, 2010).	✓ (implied)	✓	✓	✓	✓

[a] Some definitions of compassion explicitly include an element of 'recognising suffering', whereas in others, this is implied. We have indicated whether 'recognising suffering' is explicitly stated or implied in the following way: ✓(explicitly stated) and ✓ (implied).

Compassion: A Proposed Definition

To bring together the various definitions and considerations above and to aid the review of existing measures of compassion, we propose a new definition of compassion as a cognitive, affective, and behavioural process consisting of the following five elements that refer to both self- and other-compassion: 1) recognising suffering, 2) understanding the universality of suffering in human experience; 3) feeling for the person suffering and connecting with the distress (emotional resonance), 4) tolerating uncomfortable feelings aroused in response to the suffering person (e.g., distress, anger, fear) so remaining open to and accepting of the person suffering, and 5) motivation to act/acting to alleviate suffering.

We use this proposed new definition of compassion to organise the remainder of this paper, which provides a systematic review of self- and observer-rated measures of compassion. The psychometric properties of identified measures are reported and rated for quality, including the extent to which they measure each of the five elements outlined above.

Method

Inclusion and Exclusion Criteria

To be included in the main review, measures had to: 1) be available in English, 2) include a scale explicitly defined by its authors as measuring compassion, 3) include a psychometric paper outlining the development of the scale, and 4) be obtainable either within a published article or from the author (two attempts were made to contact the relevant authors to obtain measures where necessary). Measures were excluded if they did not assess participants' levels of compassion per se (e.g., measures of barriers to feeling compassion, fear of compassion, and empathy were excluded), used non-questionnaire measures of compassion, or included only a subscale on compassion. Because we do not yet know the relationship between compassion for others and self-compassion, measures of self-compassion were included because many conceptualisations and definitions of compassion do not distinguish between other- and self-compassion.

Information Sources

The databases searched for relevant measures included Web of Science (Thomson Reuters), PsycInfo, and Medline, from inception to 23 September 2015. Dissertations and theses that met the inclusion criteria were reviewed along with papers published in peer-reviewed journals. Where relevant, the most recent versions of measures were reviewed.

Search Strategy

All articles including the word 'compassion*' in combination with 'measure*', 'scale*', 'instrument*', or 'questionnaire*' in either the title, abstract, or key words were identified. Where identified papers referred to additional scales, reference lists were searched and any additional relevant papers retrieved. Experts in the field were also consulted to ensure that no measures were missed.

Assessment of Quality

The psychometric properties of each measure were reviewed and measures were rated for quality based largely on Terwee et al.'s (2007) quality criteria for health status measures. These criteria were used because they include explicit criteria for what constitutes good measurement properties. However, since these criteria relate to measures of health status, Barker, Pistrang, and Elliott's (2002) 'rules of thumb' for evaluating psychological measures were also drawn on where these seemed more appropriate[1]. Terwee et al. award measures a positive (+), intermediate (?), or negative (-) rating, or a rating of 0 where no information regarding the relevant criteria is provided. In this review, in order to make scores easier to interpret, measures were given a score of 2 if there was evidence for the criterion being fully met, 1 if the criterion was partially met, and 0 if the criterion was not met, or if no relevant data were reported. Scores were aggregated to provide an overall rating. Two researchers

[1] Terwee et al. (2007) proposed the following eight quality criteria to evaluate health status measures: 1) content validity, 2) internal consistency, 3) criterion validity, 4) construct validity (convergent and discriminant validity), 5) reproducibility (test-retest reliability), 6) responsiveness, 7) floor and ceiling effects, and 8) interpretability. We did not include criterion validity and responsiveness as criteria, for the reasons stated in the paper. Terwee et al. did not provide rules of thumb in terms of the size of correlation coefficients for the test-retest reliability criterion. They also did not account for the size of correlations for the convergent and discriminant validity criterion. Therefore, for these two criteria, we drew on Barker at al.'s (2002) general recommendations when evaluating the reliability and validity of psychological measures. We also included factor structure as a criterion.

independently scored the measures using these criteria, and any discrepancies in scoring were resolved collectively. Specifically, measures were rated across the following domains:

1. Content validity (the extent to which the domain of interest was comprehensively sampled by the items in the questionnaire). In this case, the domain of interest was considered to be compassion as defined in this review, rather than as defined by the scale's authors. Under this criterion, Terwee et al. (2007) also emphasise the importance of both members of the target population and experts being involved in item development. For this criterion to be fully met, all five elements of compassion must be captured by the items and items must have been generated in consultation with experts and members of the intended population.
2. Factor structure (whether or not the factor structure for the measure has been examined and supported). This criterion was included in addition to those proposed by Terwee et al. (2007). This criterion was scored as follows. A score of 2 was given where exploratory factor analysis (EFA) followed by confirmatory factor analysis (CFA) have been conducted in independent samples OR where CFA has been conducted if the factor structure has been previously proposed theoretically (a score of 2 was only given if the factor analyses support the proposed factor structure). A score of 1 was given if only EFA has been conducted (without CFA) and if the EFA supports the factor structure. A score of 0 was given where either factor analysis has not been conducted OR where EFA and/or CFA have been conducted that do not support the proposed factor structure.
3. Internal consistency (the extent to which items in a (sub)scale are intercorrelated and thus measuring the same construct). For this criterion to be fully met, in line with Terwee et al.'s (2007) criteria, factor analyses had to have been performed on an adequate sample size (7 * number of items and $N > 100$) and Cronbach's alpha for each identified factor had to be between .70 and .95.
4. Test-retest reliability. Based on Barker et al.'s (2002) 'rules of thumb', test-retest reliabilities had to be at least $r = .70$ for this criterion to be fully met.
5. Convergent and discriminant validity (the extent to which scores on a particular scale relate to other measures in a manner consistent with theoretically derived hypotheses). For this criterion to be met, Terwee et al. (2007) require that 1) specific hypotheses are formulated by the scale's authors about expected correlations and 2) at least three quarters of results are in line with expectations. As Terwee et al. do not

take into account the strength of these correlations, we also drew on Barker et al. (2002) and required that at least two correlations with theoretically related constructs were at least $r = .50$ to demonstrate convergent validity.

6. Floor and ceiling effects (the number of respondents achieving the highest or lowest possible scores). This was rated based on Terwee et al.'s (2007) criterion that no more than 15% of the sample should receive the top or bottom score on a scale.
7. Interpretability (how differences in scores on the measure can be interpreted, or the degree to which qualitative meaning can be attached to quantitative scores). Terwee et al. (2007) require means and SDs of scores from at least four relevant subgroups of participants to be reported (e.g., compassion scores in males vs. females, meditators vs. non-meditators) and minimal important change defined. However, as minimal important change was arguably not entirely relevant to the measures in this review, consideration was instead given to whether the authors indicated how scale scores might be interpreted.

Terwee et al.'s (2007) quality ratings also include 'criterion validity' (the extent to which scores on a particular scale relate to a 'gold standard') and 'responsiveness' (the ability of a scale to detect change over time). However, these two criteria were not rated. In the case of 'criterion validity' this was because there is no gold standard compassion measure to rate scales against. In the case of 'responsiveness' it was because such data were not typically available and as this criterion relates to clinically meaningful change, arguably the majority of the scales were not primarily designed to measure this. Therefore, the total possible score for any measure was 14.

Results

Review of Identified Measures

Figure 9 shows a flow diagram illustrating the search process. After removing duplicates, 2,146 papers were identified, with only nine measures included after screening titles, abstracts, and full texts. Table 13 provides the quality ratings of the reviewed measures and Tables 14 and 15 outline the psychometric properties of the measures. Floor and ceiling effects are not included in Tables 14 and 15 because none of the studies reported on them. Similarly, although most studies included measures of

related constructs to test convergent validity, none included measures of theoretically unrelated constructs; therefore, discriminant validity is not included in Table 15.

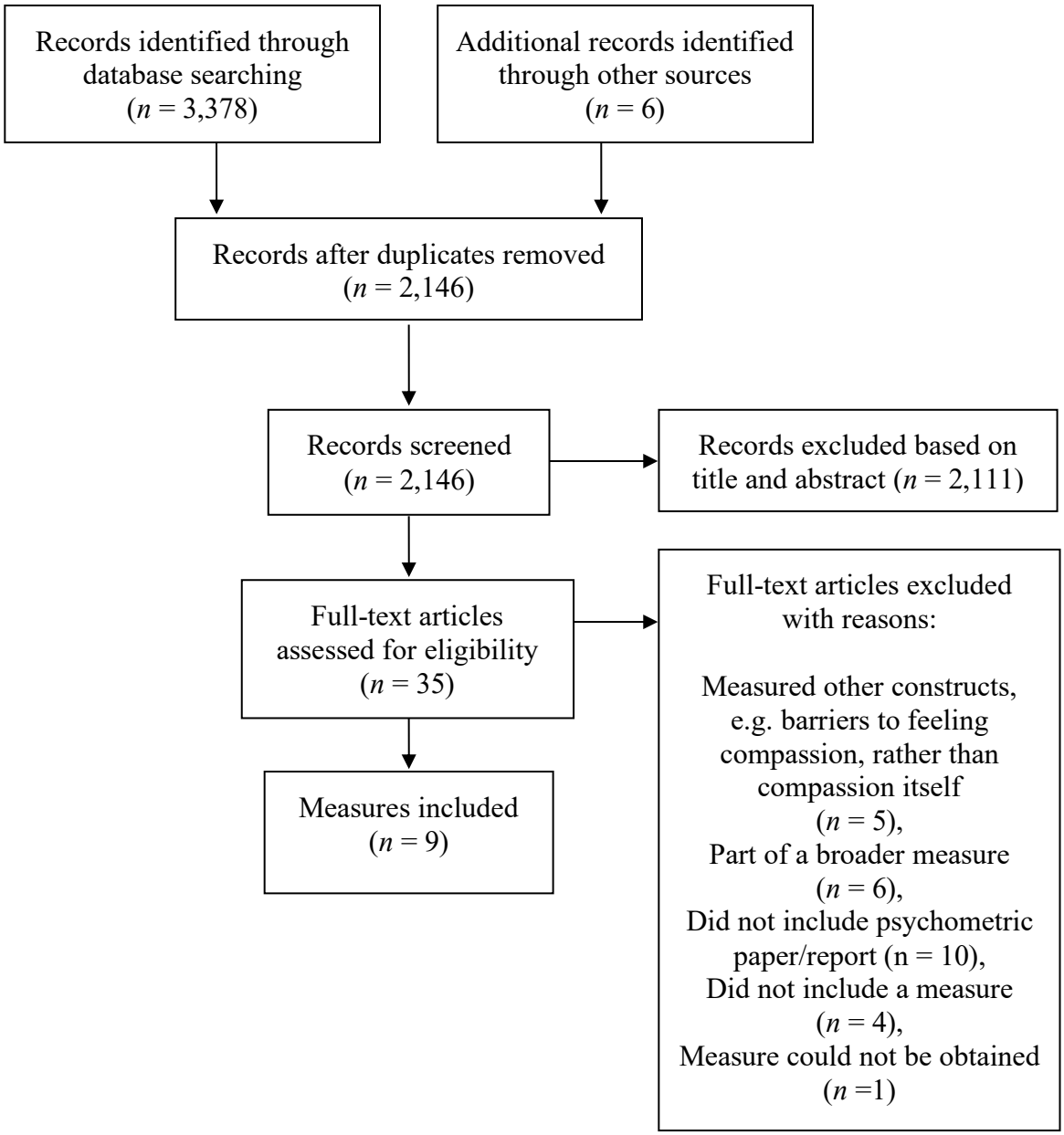


Figure 9. PRISMA flow diagram of search strategy.

Table 13.

Quality ratings of measures of compassion

Measure	Content validity	Factor structure	Internal consist-ency	Test retest reliability	Convergent / discriminant validity	Floor/ ceiling effects	Interpret -ability	**Total /14**
SCS	1	0	2	2	1	0	1	**7**
RCS	1	2	2	0	1	0	1	**7**
CLS	1	1	2	0	1	0	1	**6**
SCBCS	1	1	2	0	1	0	1	**6**
CS-P	1	2	1	0	1	0	1	**6**
CS-M	1	0	2	0	0	0	1	**4**
CCAT	1	0	2	0	0	0	1	**4**
SCS-SF	1	2	1	0	0	0	0	**4**
SCCCS	1	0	0	0	1	0	0	**2**

Note. Rating: 0 = criterion not met/insufficient data to rate criterion; 1 = criterion partially met; 2 = criterion fully met.

CCAT = Compassionate Care Assessment Tool; CLS = Compassionate Love Scale; CS-P = Pommier Compassion Scale; CS-M = Martins et al. Compassion Scale; RCS = Relational Compassion Scale; SCBCS = Santa Clara Brief Compassion Scale; SCCCS = Schwartz Center Compassionate Care Scale; SCS = Self-Compassion Scale; SCS-SF = Self-Compassion Scale – Short Form.

Table 14.

Psychometric properties of measures of compassion (content validity, factor structure, internal consistency, and test-retest reliability)

Measure	Content validity: Factors captured[a]/ 5	Content validity: Item generation (recipient and expert groups consulted?)	Proposed factor structure	Support for factor structure: Type of analysis conducted (factor structure found)[b]	Internal consistency: Adequate sample size for factor analyses?	Internal consistency: Cronbach alpha (for total scale and subscales)	Test retest reliability: *r* (time between testing)
CLS	4 (U, ER, T, & A)	Recipients = yes Experts = yes	Not reported	EFA (single factor for both versions)	Yes (N = 354)	α = .95 for both close others and strangers-humanity versions	Not reported
SCBCS	2 (ER & A)	Recipients = yes Experts = yes	Not reported	EFA (single factor)	Yes (N = 223)	α = .90	Not reported
CS-M	1 (A)	Recipients = yes Experts = yes	Five factors	EFA (two-factor structure found but rejected in favour of a single-factor model)	Yes (N = 310)	α = .82	Not reported
SCS	4 (U, ER, T, & A)	Recipients = no Experts = yes	Six factors represented	CFA (six factors)	Yes (N = 391)	Total α = .92, Subscales = .75 to .81.	Total scale: r = .93,

			under a single overarching construct				Subscales: r = .80 to .88 (3 weeks)
SCS-SF	4 (U, ER, T, & A)	Recipients = no Experts = yes	Six factors represented under a single overarching construct	CFA (six factors represented under a single overarching construct)	Yes (N = 415)	Total α = .86, Subscales = .54 to .75.	Not reported
CS-P	4 (R, U, ER, & A)	Recipients = no Experts = yes	Six factors represented under a single overarching construct	CFA (six factors represented under a single overarching construct)	Yes (sample 1: N = 439, sample 2: N = 510)	Total α = .90 (sample 1) and .87 (sample 2). Subscale αs < .70 for 4/6 subscales in sample 1 and 1/6 subscales in sample 2.	Not reported
RCS (compassion for others	4 (R, ER, T, & A)	Recipients = no Experts = yes	Four factors	CFA (four factors)	Yes (N = 231)	Subscales = .74 to .84	Not reported

and self-compassion subscales)							
CCAT	3 (ER, T, & A)	Recipients = yes Experts = yes	Not reported	EFA (four factors)	Yes (N = 250)	Total α > .70 (exact value not given), Subscales = .77 to .87.	Not reported
SCCCS	3 (R, ER, A)	Recipients = yes Experts = yes	Not reported	EFA (single factor but analysis was not conducted on all items) CFA (single factor but analysis was not conducted on all items)	Yes (N = 801)	Subscales = .97 and .95. Single scale but total α missing.	Not reported

CCAT = Compassionate Care Assessment Tool; CFA = confirmatory factor analysis; CLS = Compassionate Love Scale; CS-P = Pommier Compassion Scale; CS-M = Martins et al. Compassion Scale; EFA = exploratory factor analysis; RCS = Relational Compassion Scale; SCBCS = Santa Clara Brief Compassion Scale; SCCCS = Schwartz Center Compassionate Care Scale; SCS = Self-Compassion Scale; SCS-SF = Self-Compassion Scale – Short Form.

[a] Five elements of compassion: R = recognising suffering; U = understanding the universality of suffering; ER = emotional resonance; T = tolerating uncomfortable feelings; A = acting or motivation to act to alleviate suffering.

[b] For details of the factors identified, refer to the results section.

Table 15.

Psychometric properties of measures of compassion (convergent validity and interpretability)

Measure	Convergent validity: Correlation (Pearson's *r*) of compassion measure with measures of related constructs	Interpretability: Subgroups tested for differences
CLS	PSP other-oriented empathy subscale and empathy items from Schieman & Van Gundy (2000): *r* = .50 to .68; PSP helpfulness subscale: *r* = .23 (close others), *r* = .32 (strangers); Frequency of church attendance: *r* = .22 (close others), *r* = .43 (strangers); Volunteerism items from Mikulincer et al. (2005): *r* = .18 (close others), *r* = .35 (strangers); Social support (developed by authors): *r* = .27 (strangers), *r* = .51 (close others); DSES: *r* = .39 (close others), *r* = .44 (strangers).	Gender (women scored significantly higher than men on both versions)
SCBCS	IRI empathic concern subscale: *r* = .65; VIQ: *r* = .48; SCSORF: *r* = .27.	Gender (women scored significantly higher than men)
CS-M	CLS: *r* = .66	Gender (women scored significantly higher than men), Education (significantly higher for those with University education versus High School or less),

		Income (significantly higher for those with annual income $40,000+ versus $10,000 or less), Age, race, and marital status (no differences)
SCS	DEQ self-criticism subscale: *r* = -.65; SOC: *r* = .41; TMMS attention subscale: *r* = .11; TMMS clarity subscale: *r* = .43; TMMS repair subscale: *r* = .55; RSES: *r* = .59.	Gender (women scored significantly lower than men)
SCS-SF	Not reported	Not reported
CS-P	SCS: *r* = .01; SOC: *r* = .41; 3D-WS reflective subscale: *r* = .26; 3D-WS cognitive subscale: *r* = .39; 3D-WS affective subscale: *r* = .56; QMEE: *r* = .58; IRI empathic concern subscale: *r* = .65; IRI perspective taking subscale: *r* = .35; CLS close others version: *r* = .54; CLS strangers version: *r* = .27; SMQ: *r* = -.12.	Gender (women scored significantly higher than men)
RCS (other-compassion and self-	RCS compassion for others subscale & SCS: *r* = .24; RCS self-compassion subscale & SCS: *r* = .65; RCS compassion for others subscale & EACS emotional expression & processing subscales: *r* = .41 and .42; RCS self-compassion subscale & EACS emotional expression & processing subscales: *r* = .51 and .46; RCS compassion for others subscale & SCSRS inadequate &	Significant differences in RCS scores between Arts and Engineering students. The direction of the results for each subscale was not specified.

compassion subscales)	hated self subscales: *r* = .03 and .12; RCS self-compassion subscale & SCRS inadequate & hated self subscales: *r* = -.29 and -.41; RCS compassion for others subscale & SCSRS reassured self subscale: *r* = .01; RCS self-compassion subscale & SCRS reassured self subscale: *r* = .43; RCS compassion for others subscale and RSQ secure attachment: *r* = .34; RCS self-compassion subscale and RSQ secure attachment: *r* = .31; RCS compassion for others subscale and RSQ insecure attachment styles (fearful, preoccupied, dismissing, anxious, and avoidant): *r* = -.23, -.06, -.15, -.19, and -.22, respectively; RCS self-compassion subscale and RSQ insecure attachment styles (fearful, preoccupied, dismissing, anxious, and avoidant): *r* = -.22, -.15, -.05, -.03, and -.07, respectively.	
CCAT	Not reported	Gender (women scored carers significantly higher than men), Marital status and reason for hospitalisation (no differences)
SCCCS	Overall satisfaction with recent hospitalisation (item set 1): *r* = .54; Overall satisfaction with recent hospitalisation (item set 2): *r* = .60; Satisfaction with communication and emotional support (item set 1): *r* = .72; Satisfaction with communication and emotional support (item set 2): *r* = .64.	Not reported

3D-WS = 3-Dimensional Wisdom Scale (Ardelt, 2003); CCAT = Compassionate Care Assessment Tool; CLS = Compassionate Love Scale; CS-P = Pommier Compassion Scale; CS-M = Martins et al. Compassion Scale; DEQ = Depressive Experiences Questionnaire (Blatt, D'Afflitti, & Quinlan, 1976); DSES = Daily Spiritual Experience Scale (Underwood & Teresi, 2002); EACS = Emotional Approach Coping Scale (Stanton et al., 2000); IRI = Interpersonal Reactivity Index (Davis, 1980); PSP = Prosocial Personality Battery (Penner et al., 1995); QMEE = Questionnaire Measure of Emotional Empathy (Mehrabian & Epstein, 1972); RCS = Relational Compassion Scale; RSES = Rosenberg Self-Esteem Scale (Rosenberg, 1965); RSQ = Relationship Scales Questionnaire (Griffin & Bartholomew, 1994); SCBCS = Santa Clara Brief Compassion Scale; SCSORF = Santa Clara Strength of Religious Faith Questionnaire (Plante & Boccaccini, 1997); SCCCS = Schwartz Center Compassionate Care Scale; SCS = Self-Compassion Scale; SCS-SF = Self-Compassion Scale – Short Form; SCSRS = Self-Criticising/Attacking and Self-Reassuring Scale (Gilbert et al., 2004); SMQ = Southampton Mindfulness Questionnaire (Chadwick et al., 2008); SOC = Social Connectedness Scale (Lee & Robbins, 1995); TMMS = Trait Meta-Mood Scale (Salovey et al., 1995); VIQ = Vocational Identity Questionnaire (Dreher et al., 2007).

[a] Five elements of compassion: R = recognising suffering; U = understanding the universality of suffering; ER = emotional resonance; T = tolerating uncomfortable feelings; A = acting or motivation to act to alleviate suffering.

Compassionate love scale (CLS; Sprecher & Fehr, 2005). The CLS consists of 21 self-report items, rated on a Likert scale from 1 (not at all true of me) to 7 (very true of me). The CLS is intended for the general population and consists of two forms: one relating to significant others (including family members and friends) and one focusing on strangers and humanity at large.

Content validity. The scale was rated as partially satisfactory for content validity. Items were generated by the investigators based on a review of the literature on love and altruism and also based on a prototype analysis with laypeople around their concept of compassionate love. In line with our definition of compassion, the scale includes items related to four of our five elements of compassion: feeling moved by other people's suffering (emotional resonance), understanding or imagining something about their condition as a fellow being, accepting and not judging others (which implies tolerance), and being motivated to help them. However, the CLS did not appear to contain items explicitly related to recognising suffering.

Three items include the word 'compassion' or 'compassionate love', which requires respondents to define these concepts themselves. However, it seems uncertain whether they will know what is meant by 'compassionate love' or define it uniformly. Additionally, not all items on the scale relate to those who are suffering, and it is questionable whether items such as "I feel happy when I see that [loved ones/others (strangers)] are happy" and "I very much wish to be kind and good to [my friends and family members/fellow human beings]" assess compassion or empathy and kindness, respectively. Finally, the scale refers explicitly to either close others or strangers but does not allow respondents to consider people who may not fall into either of these categories (e.g., patients responding in relation to healthcare professionals), potentially limiting its use in some contexts.

Factor structure and reliability. Exploratory factor analysis (EFA) yielded a single factor structure for each version of the scale. Sprecher and Fehr (2005) did not explicitly propose a factor structure for the CLS prior to analysis and did not conduct CFA. Internal consistency was high for both versions. Test-retest reliability was not reported.

Convergent validity and interpretability. Convergent validity was supported by significant correlations in the expected directions with measures of empathy, helpfulness, volunteerism, and spiritual experiences. Limited subgroup analyses were

undertaken by Sprecher and Fehr (2005), showing that women obtained significantly higher compassion scores than men on both versions.

Santa Clara brief compassion scale (SCBCS; Hwang, Plante, & Lackey, 2008). The SCBCS is a shortened version of Sprecher and Fehr's (2005) CLS, consisting of five items from the original scale (the correlation between the two scales is $r = .95$). Unlike the CLS, this scale refers exclusively to strangers rather than to close others. The items of the SCBCS were selected because they had moderate means, high standard deviations, and high correlations with the overall score from the CLS.

Content validity. The SCBCS was rated partially satisfactory for content validity. The scale includes items related to two of our five elements of compassion: emotionally connecting with other people's suffering and acting to help them. However, unlike the CLS, the SCBCS did not appear to contain items explicitly related to understanding the universality of suffering and tolerating uncomfortable feelings, and also did not include items explicitly related to recognising suffering. Two items contain the word 'compassion', again relying on respondents to define this term rather than tapping into its underlying elements.

Factor structure and reliability. EFA yielded a single factor structure for the SCBCS and CFA was not conducted. Internal consistency was high. Test-retest reliability was not reported.

Convergent validity and interpretability. The SCBCS was strongly correlated with empathic concern, moderately correlated with vocational identity, and showed a small correlation with strength of religious faith. Examination of group differences was limited to gender and showed that women scored significantly higher than men.

The compassion scale (CS-M; Martins, Nicholas, Shaheen, Jones, & Norris, 2013). Martins et al.'s CS-M is a 10-item self-report scale developed to measure five domains of compassion: generosity, hospitality, objectivity, sensitivity, and tolerance across social networks and relationships (strangers, friends, and family) using a 1 (none) to 7 (all) response scale. The aim of the scale was to provide a measure of compassion across domains that could be enhanced through training, as the authors argue that scales like the CLS do not lend themselves well to measuring compassion in a way that can be targeted for education. Items were generated and evaluated by a panel of academic and community experts.

Content validity. Martins et al.'s (2013) scale was rated partially satisfactory for content validity. The CS-M focuses exclusively on practical acts of compassion

including giving financial help to others, using your free time to help others, and doing things for others at a cost or risk to yourself or your family and friends. Thus, only the acting to alleviate suffering factor of our five-factor definition is captured by the items of the CS-M; items related to recognising suffering, understanding the universality of suffering, emotional resonance, and tolerating uncomfortable feelings were not included. Additionally, it could be argued that the scale's items measure only a limited range of acts of compassion (giving away money, using free time to help others, sharing personal space, or doing something for others at a cost to oneself) and if the scale were applied to certain contexts (e.g., a healthcare context), the items may not assess the types of actions that might be expected in those contexts (e.g., considering ways to make those who are suffering more comfortable). Indeed, it is not altogether clear for what population the scale is intended. Furthermore, items such as "How many times would you do the right thing if it puts your family at risk?" do not appear to fit well with the response scale, which ranges from 'none' to 'all'.

Factor structure and reliability. EFA did not support the proposed five-factor structure; the analysis identified a two-factor solution. However, the two-factor structure was rejected by the authors in favour of a single factor model, arguing that, as all items beginning "How much of your…?" loaded onto one factor and all items beginning "How many times would you…?" loaded onto the second factor, the factors appeared to reflect methodological differences between items rather than substantively different constructs. Cronbach's alpha for the total scale was acceptable. Test-retest reliability was not tested.

Convergent validity and interpretability. The authors only compared their scale with the CLS ($r > .50$). In terms of interpretability, the authors provided mean scores for a range of subgroups, however, although they argue that the scale should help measure change in compassion after training, they do not provide any indication of what level of change on the scale would be needed to show that such training had been of value.

Self-compassion scale (SCS; Neff, 2003b). The SCS is a 26-item scale with a 5-point response scale from 1 (almost never) to 5 (almost always).

Content validity. The scale was rated partially satisfactory overall for content validity. Although items were selected after extensive piloting, it is notable that this was only carried out with experts and undergraduate students, even though the scale's target population included community and clinical samples. The scale includes items

related to four of the five elements in the definition of compassion used in this review: understanding the universality of suffering, emotional resonance, the ability to tolerate distressing feelings, and feeling motivated to act or acting to help ameliorate one's suffering. However, the scale does not include items specifically relating to being attentive to how one is feeling.

Factor structure and reliability. CFA of the SCS supported the six-factor model, with each of the three components of self-compassion split into two sub-factors - one comprising 'positively' worded and one 'negatively' worded items. This resulted in the following factors: kindness versus self-judgement; mindfulness versus over-identification; and common humanity versus isolation. However, Neff (2003b) found only a marginal fit with a single higher-order factor, questioning whether the six factors can be explained by a single overarching construct of self-compassion. Other studies have also questioned the higher-order factor structure and the non-hierarchical six-factor model across a range of populations, student, clinical and meditating/non-meditating (e.g., Costa, Marôco, Pinto-Gouveia, Ferreira, & Castilho, 2015; López et al., 2015; M. J. Williams et al., 2014). Several studies have suggested a two-factor model of self-compassion, with the factors representing the positive and negative dimensions of self-compassion and self-criticism, respectively (e.g., Costa et al., 2015; López et al., 2015).

Recently, Neff (2015) argued that the two-factor conceptualisation of the SCS is problematic in that it does not capture the relative balance between the three proposed broad components of self-compassion (self-kindness vs. self-judgment, common humanity vs. isolation, and mindfulness vs. over-identification). Instead, Neff proposed a bifactor model of self-compassion, where each item loads directly onto a general factor as well as their respective subscale, and suggests that researchers can select whether to analyse subscale scores separately or use a total SCS score depending on their interests.

Cronbach's alpha values for total SCS scale and subscale scores and test-retest reliability were adequate.

Convergent validity and interpretability. Convergent validity was supported by significant correlations in the expected direction between the SCS and other related measures, several of which were ≥ .50. A partially satisfactory score was achieved for interpretability, as only gender differences were reported.

Self-compassion scale: Short form (SCS-SF; Raes, Pommier, Neff, & Van Gucht, 2011). Raes et al. developed a 12-item version of the SCS by selecting two items from each of the SCS's six subscales, based on their high correlations with the SCS and intended subscales, and high intercorrelations. The SCS-SF is rated in the same way as the SCS.

Content validity. The scale was rated partially satisfactory for content validity for the same reasons as the long form.

Factor structure and reliability. CFA supported the proposed six-factor hierarchical structure of the measure. Internal consistency was acceptable for the total score but was variable for the individual subscales. Test-retest reliability was not reported.

Convergent validity and interpretability. Relevant data were not reported for convergent validity and interpretability.

The compassion scale (CS-P; Pommier, 2010). The CS-P is a 24-item self-report scale targeted at the general population and based on the argument that compassion consists of six elements: kindness (in contrast to indifference), mindfulness (in contrast to disengagement), and common humanity (in contrast to separation). Responses are given on a five-point Likert scale, ranging from 1 (almost never) to 5 (almost always).

Content validity. The CS-P was rated partially satisfactory for content validity. Items were devised by the author, based on theory and research, and reviewed by a panel of experts. The scale includes items consistent with four of our five elements of compassion: recognising suffering, feeling moved by suffering, understanding or imagining something about another person's condition as a fellow being, and motivation to act/acting to alleviate suffering. Although in her development of the scale Pommier (2010) notes that compassion requires the ability to tolerate uncomfortable feelings in the face of suffering so that one can remain tolerant and accepting of others, the scale does not appear to directly assess this, other than asking whether respondents "try to keep a balanced perspective on the situation" when people tell them about their problems, or whether they tend to avoid those who suffer.

Additionally, several of the scale's items include the words 'sometimes', 'often', or 'usually' which conflict with the response scale used ('almost never' to 'almost always') and makes responses difficult to interpret. The response scale is perhaps also unintuitive for negatively worded items – for example, a response of "I

almost never don't feel emotionally connected to people in pain" may be difficult for some people to rate accurately. Similarly, items such as "Suffering is just a part of the common human experience" cannot be answered accurately using the scale from 'almost never' to 'almost always' and do not sit well with the scale's instructions to "indicate how often you behave in the stated manner".

Factor structure and reliability. CFA supported the proposed six-factor structure of the measure and that a single higher order factor of compassion explained the intercorrelations between the six factors. EFA was not conducted because Neff's (2003b) SCS had already demonstrated these six factors. However, as noted earlier, compassion for others and self-compassion were not significantly correlated in Pommier's (2010) research, suggesting that the factor structure for each measure cannot be assumed to be identical. Internal consistency was high for the total score but mixed and inconsistent across samples for the subscales. Test-retest reliability was not reported.

Convergent validity and interpretability. Convergent validity was supported by significant correlations in the expected direction between the CS-P and other measures of compassion, empathy, perspective-taking, and wisdom; several of these were ≥ .50. However, the CS-P was not significantly correlated with the SCS (Neff, 2003b), a problematic finding for the scale's construct validity, given that the CS-P was developed based on the factor structure of the SCS. Additionally, while the scale was positively correlated with the CLS, this correlation was small for the strangers-humanity version (r = .27; r = .54 for the close others version). This suggests that the CS-P and the CLS may not be measuring the same construct. Another unexpected finding was that the Southampton Mindfulness Questionnaire (SMQ; Chadwick et al., 2008) had a small negative correlation with the CS-P. The only subgroup analysed was gender, again showing that women scored higher than men.

Relational compassion scale (RCS; Hacker, 2008). The RCS consists of 16 items rated on a four-point scale (from 'do not agree' to 'agree strongly'). The scale consists of four subscales which measure respondents' compassion for others and self-compassion, along with their beliefs about how compassionate other people are to each other, and their beliefs about how compassionate other people are to them. The latter two subscales extend beyond simply measuring respondents' own levels of compassion, but the scale was nevertheless included because it defines itself as a

comprehensive measure of compassion and also because the subscales were psychometrically tested individually.

Content validity. The RCS was rated partially satisfactory for content validity. The scale's items that comprise the 'compassion for others' subscale assess people's capacity to recognise and understand suffering and accept and not judge others (which implies tolerance), just two of the five elements in our definition of compassion. Additionally, some items relate to other people's experiences in general, rather than specifically to their suffering. The items comprising the self-compassion subscale assess emotional resonance and acting to alleviate suffering, two of the five elements in our definition. Items related to understanding the universality of suffering were not included in either subscale.

Factor structure and reliability. CFA using the final version of the RCS supported its proposed four-factor structure. Internal consistency was acceptable for all four subscales. Test-retest reliability was not tested.

Convergent validity and interpretability. Although several correlations with related measures were ≥ .50, specific hypotheses appear not to have been set out in advance about the expected direction of correlations and in the discussion, the author highlights some unexpected findings. For example, the 'compassion for others' subscale did not correlate significantly with a measure of self-criticism/self-attack and self-reassurance (Gilbert, Clarke, Hempel, Miles, & Irons, 2004). Only one subgroup was analysed for interpretability (Arts versus Engineering students), however the authors stated no predictions about differences between these groups.

Compassionate care assessment tool (CCAT; Burnell & Agan, 2013). The 28-item CCAT was developed to measure levels of compassion demonstrated by individual nurses providing care for patients in acute hospital settings. In contrast to the other scales reviewed so far, this scale is completed by patients in relation to their carers. Respondents rate compassionate care from two perspectives - the importance of each item to them personally, and the degree to which their individual nurses demonstrated these qualities. Ratings range from 1 to 4. A selection of possible items for the scale were derived from the Spiritual Needs Survey (Galek, Flannelly, Vane, & Galek, 2005) and the Caring Behaviours Inventory (Wu, Larrabee, & Putman, 2006), and refined after consulting with hospital staff involved in implementing national criteria for compassionate care, nurses, and patients. The CCAT focuses on four domains: the ability of carers to establish meaningful connection (e.g., having a sense

of humour), to meet patient expectations (e.g., giving timely treatments), display caring attributes (e.g., considering personal needs), and exhibit capable practitioner qualities (e.g., appearing competent).

Content validity. Overall, the scale was considered partially satisfactory for content validity. It includes items relating to three of the five elements in our definition of compassion: whether patients thought carers felt for them (emotional resonance), acted to help relieve their suffering, and could tolerate distress (e.g., asking if they "remained calm at all times", treated them non-judgmentally, and excused their shortcomings). Items related to recognising suffering and understanding the universality of suffering were not included. Additionally, some of the items are rather ambiguous - for example, one item asks whether nurses "addressed difficult issues", which could relate to their ability to tolerate distress, or to their ability to resolve more practical matters. It is also questionable whether the scale is actually measuring levels of compassion of nurses; factor analyses appear to have been carried out based on asking patients to rate how important each item was to them, rather than on asking them to rate the extent to which their carers behaved in this way.

Furthermore, as a number of items were derived from the Spiritual Needs Survey, there is a fairly strong emphasis on whether spiritual support was offered to patients, which is not necessarily relevant to the measurement of compassion for all patients. Similarly, several items taken from the Caring Behaviours Inventory ask about whether nurses gave timely treatments to patients, showed skill with equipment, and helped control pain; however, while competence may be important in order to provide compassionate care, such abilities in themselves do not necessarily equate to showing compassion. It could also be argued that some of the areas tapped, such as providing timely treatments, controlling pain, and providing access to spiritual support, depend on variables outside of nurses' power (i.e., managerial or organisational level factors), and this raises a wider issue around the extent to which compassion can and should be measured at an individual or organisational level.

Factor structure and reliability. EFA supported a four-factor structure, with the four aforementioned domains. However, as previously noted, analyses appear to have been carried out based on asking patients to rate how important each item was to them, rather than on asking them to rate the extent to which their carers behaved in this way. This means that it is not clear whether the scale is measuring actual levels of compassion of their nurses per se. Additionally, the authors report that only 20 of the

28 items fit into the four factors identified, but they nonetheless appear to have retained all 28 items. Therefore, the CCAT was given a rating of 0 for factor structure. Cronbach's alpha values were adequate for the total scale and subscales. Test-retest reliability was not reported.

Convergent validity and interpretability. Convergent validity was not reported. Limited subgroup analyses were conducted for interpretability.

The Schwartz Center compassionate care scale (SCCCS; Lown, Muncer, & Chadwick, 2015). The 12-item SCCCS was developed to measure patients' ratings of compassionate inpatient care received from physicians' during a recent hospitalisation. Patients complete items using a 10-point scale from 1 (not at all successful) to 10 (very successful). Items were initially developed by a committee consisting of patients, family members of patients, and individuals working in healthcare policy and advocacy, and were fine-tuned in five focus groups with patients, physicians, and nurses.

Content validity. Overall, the SCCCS was considered partially satisfactory for content validity. It includes items which could be interpreted to relate to three of the five elements in our definition of compassion: whether patients thought physicians expressed sensitivity, care, and compassion for them (emotional resonance/acting to alleviate suffering), listened attentively (recognising suffering), and acted in ways to relieve their suffering. The SCCCS did not appear to contain items related to understanding the universality of suffering and tolerating uncomfortable feelings. Additionally, a couple of items refer to competence in caring (whether physicians spend enough time with patients, whether physicians communicate test results in a timely manner) which does not necessarily equate to showing compassion and could be dependent on factors outside of physicians' power (i.e., managerial or organisational level factors).

Factor structure and reliability. The SCCCS originally consisted of 16 items which were split into two item sets and administered to 801 recently hospitalised patients; half were asked item set one and half item set two. The authors conducted an EFA and CFA for each set of items and concluded that items within each set were unidimensional. However, they did not conduct analyses on all of the items, making it impossible to determine whether the scale as a whole is unidimensional, or whether the measure consisted of two separate scales or subscales. Despite this, the SCCCS was presented as a single scale. Although Cronbach's alpha values were adequate for both

sets, these values were based on there being eight items in each set; the final 12-item scale consisted of seven items from the first set and five items from the second set after the removal of problematic items (e.g., items with lowest item-total correlations). The alpha value for all of the scale items was also missing. Test-retest reliability was not tested.

Convergent validity and interpretability. The authors found moderate-to-large, positive correlations between scores on both sets of items from the 12-item SCCCS and related constructs. Interpretability was not tested.

Discussion

The first aim of this paper was to synthesise existing conceptualisations of compassion and to propose a new definition that integrates common elements. A range of definitions from Buddhist and Western psychological perspectives were considered and five components of compassion were identified: recognition of suffering, understanding its universality, feeling for those who are suffering, tolerating the distress associated with the witnessing of suffering, and motivation to act or acting to alleviate the suffering. Each of these components has been articulated by several published definitions of compassion, although no single existing definition explicitly includes all five of them. We do not claim that these five elements constitute statistically distinct factors of an overarching construct of compassion; this possibility must be empirically tested. However, we argue that our definition provides a useful foundation for the development of a comprehensive new measure of compassion.

The need for a new measure is supported by the findings of our review of existing measures of compassion. The maximum quality rating of any measure was seven out of a possible fourteen, suggesting that no scale exists that comprehensively measures compassion and provides scores with acceptable levels of reliability and validity. In other words, we cannot be confident that existing measures of compassion are measuring this construct accurately and this raises significant barriers to scientific progress in the field – how can we assess compassion and evaluate the effectiveness of interventions intended to enhance compassion if we cannot measure the construct accurately?

Quality ratings were low both because of poor ratings for content validity (the extent to which items appeared to fit our definition of compassion) and because of poor

or untested psychometric properties. Internal consistency was strong for total scores but weak for many subscales. Evidence for the proposed factor structure of some scales was weak or absent. The presence of floor or ceiling effects was not examined for any scale, and test-retest reliability was examined for only one. Convergent correlations were generally significant and in the expected directions, but discriminant validity was not assessed. Low quality ratings could also be attributed to measures being in their early stages of development and initial papers being unlikely to include a thorough test of psychometric properties. Quality ratings for compassion measures may improve over time with additional research including psychometric research.

The strongest measures identified were Neff's (2003b) Self-Compassion Scale and Hacker's (2008) Relational Compassion Scale, but neither of these measures capture each of the five elements in our definition. As Neff's measure focuses on self-compassion rather than compassion more generally or compassion for others, it is in any case not entirely suitable as a measure that can be used to determine levels of compassion in populations for whom compassion toward others is of interest (e.g., healthcare professionals). Given the current enthusiasm for compassion across different contexts, it is critical for future research to develop a psychometrically robust measure of the proposed definition of compassion as well as to explore more fully the relationship between self- and other-compassion.

Strengths and Limitations

A strength of this review is its contribution to greater clarity in the conceptualisation of compassion and its components, which have previously been described in a variety of ways. The five elements of compassion extracted from our synthesis of definitions suggests that compassion is a complex construct that includes emotion but is more than an emotion, as it also includes perceptiveness or sensitivity to suffering, understanding of its universality, acceptance, non-judgment and distress tolerance, and acting or intentions to act in helpful ways. This conceptualisation suggests that compassion can be state-like and trait-like. Sensitivity to one's own or others' suffering, emotional responsiveness, acceptance and nonjudgment in the face of suffering, and motivation to be helpful are all likely to fluctuate across time and situations. On the other hand, Goetz et al. (2010) present evidence suggesting that compassion can be seen as a trait-like quality that endures over time (e.g., Eisenberg et

al., 2002). An implicit assumption of compassion-focused interventions seems to be that a trait-like general tendency to be compassionate toward oneself or others can be developed through repeated practice of skills that cultivate compassionate states, attitudes, or behaviours. Additionally, although many of the questionnaires reviewed treat compassion as a disposition that is fairly consistent across contexts, some measures conceptualise compassion as operating within a particular context or social interaction (e.g., the CCAT).

This review assumes that compassion can indeed be measured with questionnaire methods. Some authors have suggested that subtle but observable behaviours, such as using a soft tone of voice, may also be valid indicators of compassion (Cameron, Mazer, DeLuca, Mohile, & Epstein, 2015), while Pearson (2006) notes that acts of compassion are often invisible, being "simple not clever; basic not exquisite; peripheral not central" (p. 22). This means that, as Dewar, Pullin, and Tocheris (2011) note, "there is a danger, therefore, of measuring what is easy to quantify, rather than what is important" (p. 32). Dewar et al. also point out that compassion can be seen as something that is negotiated between individuals in their interactions. These points suggest that, as with many psychological variables, questionnaire measures may only provide a partial picture of compassion. Furthermore, while questionnaire measures benefit from being simple to administer and complete, and helpful for tapping people's underlying attitudes where these are not directly observable, it may be difficult for people to complete such measures accurately in some contexts, for example in situations where healthcare staff feel under threat to be seen to be compassionate.

A further limitation of this review is the approach taken to identifying the definitions of compassion in Table 12. A systematic search was attempted but the way the field has evolved does not easily lend itself to a systematic review. Defining compassion was very rarely the primary purpose of papers; definitions were typically embedded as secondary to addressing the primary purpose of the paper. An early search generated an unfeasibly large number of results. We therefore relied on the expertise of the authors to identify key theorists and sources in the field.

The review also assumed that individual levels of compassion *should* be measured. However, it has been argued that measuring compassion at the individual level opens people to accusations that they are not sufficiently compassionate. For example, in a healthcare context, this may result in a tendency to blame healthcare

professionals for failings that in fact relate to external factors such as resourcing pressures or organisational restructuring (Crawford et al., 2014). This is an important consideration and highlights the need to ensure that efforts to measure levels of compassion among individuals do not serve to overstate individual deficits whilst deflecting attention from the broader impact of resourcing constraints and wider organisational changes.

Future Research

This review has argued that currently no psychometrically robust self- or observer-rated measure of compassion exists, despite widespread interest in measuring and enhancing compassion towards self and others. Future research should therefore focus on developing a psychometrically robust questionnaire-based measure of compassion, whilst keeping in mind the complexities around measuring this construct. It will subsequently be of value for future research to identify interventions (at both an individual level and organisational level) that have the potential to enhance compassion and examine whether changes in compassion mediate the outcomes of these interventions.

Although our review provides a foundation for progressing research into compassion, it represents a starting point. Future work should articulate theory-driven hypotheses that test the relationships between key constructs and the validity of our five-element definition of compassion. This will generate important new knowledge about how these different elements interact to give rise to compassion. It may be that some elements are facilitators of compassion or emergent factors rather than defining features.

Using a range of designs (including prospective and experimental designs), and triangulation of measurement to include behavioural (e.g., observable compassionate responses), bio-behavioural (e.g., as derived from Gilbert's theory), and self-report measures, will further aid the development of theory and understanding. It is likely that this will have real practical implications for how best to cultivate compassion in ways that support resilience and wellbeing at both personal and societal levels.

Conclusion

In recent years, compassion has received increased scientific interest. Compassion has been defined here, in line with the literature, as involving five elements: recognising suffering in others, understanding the common humanity of this suffering, feeling emotionally connected with the person who is suffering, tolerating difficult feelings that may arise, and acting or being motivated to act to help the person. A systematic search of measures of compassion was undertaken but all of the identified measures were found to have notable psychometric weaknesses. This is a serious limitation in the field. For example, without adequate measures, we cannot determine with any confidence levels of compassion or whether interventions designed to enhance compassion are effective. Therefore, we now call for empirical testing of our five-element definition and the development of a measure of compassion, following good practice guidelines to generate items and test its psychometric properties.

Chapter 6:
An Empirical Examination of the Factor Structure of Compassion

Chapter 6 is published in *PloS one* as:

Gu, J., Cavanagh, K., Baer, R., & Strauss, C. (2017). An empirical examination of the factor structure of compassion. *PloS one, 12*(2), e0172471.

Author Contributions

Design:	JG, CS, and KC
Data Collection:	JG
Data Analysis:	JG
Initial Draft:	JG
Reviewing and Editing:	All authors

Abstract

Compassion has long been regarded as a core part of our humanity by contemplative traditions, and in recent years, it has received growing research interest. Following a recent review of existing conceptualisations, compassion has been defined as consisting of the following five elements: 1) recognising suffering, 2) understanding the universality of suffering in human experience, 3) feeling moved by the person suffering and emotionally connecting with their distress, 4) tolerating uncomfortable feelings aroused (e.g., fear, distress) so that we remain open to and accepting of the person suffering, and 5) acting or being motivated to act to alleviate suffering. As a prerequisite to developing a high-quality compassion measure and furthering research in this field, the current study empirically investigated the factor structure of the five-element definition using a combination of existing and newly generated self-report items. This study consisted of three stages: a systematic consultation with experts to review items from existing self-report measures of compassion and generate additional items (Stage 1), exploratory factor analysis of items gathered from Stage 1 to identify the underlying structure of compassion (Stage 2), and confirmatory factor analysis to validate the identified factor structure (Stage 3). Findings showed preliminary empirical support for a five-factor structure of compassion consistent with the five-element definition. However, findings indicated that the 'tolerating' factor may be problematic and not a core aspect of compassion. This possibility requires further empirical testing. Limitations with items from included measures lead us to recommend against using these items collectively to assess compassion. Instead, we call for the development of a new self-report measure of compassion, using the five-element definition to guide item generation. We recommend including newly-generated 'tolerating' items in the initial item pool, to determine whether or not factor-level issues are resolved once item-level issues are addressed.

Introduction

Until recently, the scientific study of compassion towards others has been hampered by a lack of definitional and measurement clarity (Strauss et al., 2016). This contrasts to the long-standing emphasis on compassion as a fundamental part of our humanity in Eastern contemplative traditions, such as Buddhism, and major world religions (Gilbert, 2010; Moses, 2002). In Buddhism, compassion, conceptualised as the "heart that trembles in the face of suffering" (Feldman & Kuyken, 2011), is regarded as essential to gaining wisdom (Wallace, 2001), happiness (Dalai Lama, 2002a), and freeing our minds from destructive emotions (Goleman, 2003). In support of this, emerging research has linked compassion to a number of positive constructs, such as increased happiness and self-esteem (Mongrain, Chin, & Shapira, 2011), increased social connectedness, greater wellbeing, and lowered levels of loneliness (Crocker & Canevello, 2008). Additionally, the 2015 *World Happiness Report* (Davidson & Schuyler, 2015) highlighted compassion-related constructs, altruism and prosocial behaviour, as one of the four constituents of wellbeing.

Over the past few decades, scientific and societal interest in compassion has blossomed in many different sectors. There is a growing awareness of the importance of placing compassion at the heart of healthcare for the benefit of patients, staff, and healthcare organisations. In the US, compassion is integral to the Institute of Medicine's definition of patient-centred care (Institute of Medicine, 2001) and according to a US survey, 81% of 800 service users and 71% of 510 physicians agreed that compassionate care has an impact on whether a patient lives or dies (Lown, Rosen, & Marttila, 2011). Similarly, compassion is considered to be one of the six fundamental values in the UK National Health Service constitution (Department of Health, 2013). The growing awareness of the importance of compassion is connected to concerns about compassion wearing thin in cases of work-related burnout (Joinson, 1992), a common problem in the emotionally demanding healthcare profession. The benefits of compassion have also been recognised in the education sector, with organisations such as 'Mind with Heart' (www.mindwithheart.org), the 'Compassion in Education' foundation (www.coedfoundation.org.uk), and the 'Resilience, Wellbeing, Success' programme (www.rws.today) dedicated to equipping teachers and learners with the skills necessary to promote more compassionate learning environments.

Considering compassion as an innate capacity is also gaining traction in scientific circles. From an evolutionary perspective, compassion is thought to confer reproductive advantages through its role in the care-giving system for protecting and nurturing vulnerable offspring (e.g., de Waal, 2009; Gilbert, 2005). In humans and other higher-order primates, as the mind increased in complexity and competency, compassion towards offspring was thought to generalise to others in need (e.g., de Waal, 2009; Gilbert, 2005), partly because it enables advantageous cooperative relationships with non-kin (e.g., Keltner, 2009). These reproductive advantages mean that compassion is a desirable criterion in mate selection (e.g., Keltner, 2009). This evolutionary perspective echoes that of Darwin (1871), who considered compassion, which he termed 'sympathy', to be an instinct which confers survival advantages, noting that: "Sympathy will have been increased through natural selection; for those communities, which included the greatest number of the most sympathetic members would flourish best, and rear the greatest number of offspring" (p. 130). Interestingly, recent studies also show that people regard moral capacities (including compassion) as the mental faculty most essential to their sense of self and how they perceive the identity of others, over and above autobiographical memories, personality, and desires and preferences (Strohminger & Nichols, 2014).

Cultivating Compassion

Several interventions have the potential to enhance compassion (or outcomes related to compassion). Programmes developed to explicitly target the cultivation of compassion include Compassion Cultivation Training (Jinpa, 2010), Cognitively-Based Compassion Training (designed by Lobsang Tenzin Negi), Compassionate Mind Training (Gilbert & Irons, 2005), Mindfulness-Based Compassionate Living (van den Brink & Koster, 2015), Mindful Self-Compassion (Neff & Germer, 2013), and other compassion-based contemplative practices such as Loving Kindness Meditation. In addition, mindfulness-based interventions such as mindfulness-based cognitive therapy (Segal, Williams, & Teasdale, 2002, 2013) and mindfulness-based stress reduction (Kabat-Zinn, 1982) are examples of interventions which may implicitly raise participants' level of compassion. A number of randomised controlled trials have found these interventions to be effective for improving a broad range of psychological outcomes (e.g., Desbordes et al., 2012; Fredrickson, Cohn, Coffey, Pek, & Finkel,

2008; Hutcherson, Seppala, & Gross, 2008; Jazaieri et al., 2013, 2014; Mascaro, Rilling, Negi, & Raison, 2012; Shapiro, Schwartz, & Bonner, 1998). However, none of these studies measured compassion directly; improvements in compassion were inferred from increases in constructs thought to be related to compassion. This is a significant limitation – we do not know if enhanced compassion is the mechanism through which these interventions are having their beneficial effects. The evidence for compassion-based interventions would be considerably strengthened by measuring compassion itself; however, this has been impeded by a lack of definitional clarity.

Conceptualising Compassion

Despite the growing interest in compassion, and the development of interventions designed to cultivate compassion, scientific progress has been hindered by the number of different ways in which the construct has been conceptualised. This is further complicated by the existence of closely-related constructs such as empathy, sympathy, love, altruism, kindness, and pity, and the tendency to use these terms interchangeably. In their comprehensive review of empirical studies of compassion towards others, Goetz, Keltner, and Simon-Thomas (2010) found support for compassion as a unique, multidimensional construct which can be differentiated from related states in terms of the appraisal processes, affective experience, physiological responses, and patterns of behaviour involved when encountering others in need. Their findings support the evolutionary account and further justify the study of compassion as a construct in its own right. Neuroscientific findings of differential activation of brain regions in response to empathy and compassion training (Klimecki, Leiberg, Ricard, & Singer, 2013) lend additional support to the perspective of compassion as a distinct construct.

A closely-related construct which has received more empirical attention is compassion directed towards the self, or self-compassion (Neff, 2003a). The Buddhist perspective, which has a more nuanced view of the duality of self and others, regards the underlying processes of compassion as common to our experience regardless of the object of compassion (self or others). Moreover, self-compassion is seen as supportive of compassion towards others (Dalai Lama, 2000). However, current research indicates that the relationship between compassion for others and self-compassion is not straightforward. Neff and Pommier (2013) have failed to find a relationship between

self-compassion and other-focused constructs (other-compassion, compassion for humanity, empathic concern, and altruism) in a student sample, but found significant associations between self-compassion and other-focused constructs in samples of community adults and meditators. Using an alternative measure of compassion for others, Pommier (2010) also failed to find a significant relationship between this and self-compassion in a student sample. It is currently uncertain whether the lack of association reflects a separation between these two forms of compassion, only occurs in students or in Western cultures, or can be explained by limitations of the measures used in these studies (e.g., Strauss et al., 2016).

Strauss and colleagues (2016) reviewed the compassion literature and consolidated the range of conceptualisations of compassion into one multifaceted definition. They concluded that compassion entails five elements that apply to the self or others: 1) recognising suffering, 2) understanding the universality of suffering in human experience, 3) feeling for the person suffering and emotionally connecting with their distress, 4) tolerating any uncomfortable feelings aroused (e.g., fear, disgust, distress, anger) so that we remain accepting and open to the person in their suffering, and 5) acting or being motivated to act to alleviate the suffering. Strauss et al. also systematically reviewed existing questionnaire measures of compassion and evaluated each measure's psychometric properties. The authors concluded that none of the scales reviewed comprehensively assessed all elements of compassion and many scales had poor or inadequately tested psychometric properties. They call for the empirical examination of the proposed five elements of compassion and ultimately, the development of a psychometrically robust measure that comprehensively captures the key elements of compassion. These steps are crucial if we are to progress the various strands of compassion research discussed, including investigating how compassion relates to other constructs and evaluating of the effectiveness of interventions to cultivate compassion.

The Current Study

No studies have empirically investigated the five-element definition of compassion proposed by Strauss et al. (2016). This is an essential first step to deepening our understanding of compassion, developing a high-quality measure, and furthering research in this field. The current study used factor analyses, relying

primarily on existing self-report measures of compassion, to determine whether items from this combined pool could be identified that provide preliminary support for the proposed five-element definition. However, previous examination of these items (Strauss et al., 2016) suggested two potential problems. First, the combined pool might contain too few items to represent some of the five elements (e.g., tolerating uncomfortable feelings, understanding the universality of suffering). Second, some of the existing items may have poor content validity, as they appear to reflect related constructs, such as empathy and altruism. Therefore, sufficient new items were generated to provide a meaningful test of the proposed five-factor structure. The study consisted of three stages.

Stage 1 involved a systematic consultation with expert groups to review items from existing self-report measures of compassion and to generate additional items. Strauss et al.'s (2016) comprehensive five-element definition of compassion was used to guide the item generation and review process. Stage 1 resulted in a pool of 80 items for factor analysis: 54 items from existing compassion questionnaires and 26 items generated by expert consultants (see later for details).

Stage 2 used exploratory factor analysis (EFA) in a sample of University students to examine the factor structure of the pool of compassion items generated in Stage 1. Although Strauss et al. (2016) proposed a five-element structure of compassion based on their review of the theoretical literature, EFA was conducted to remove redundant items and to provide an empirical examination of the structure of the item pool without assuming that the five-element model would emerge. Using EFA therefore allows us to explore the factor structure of compassion naturally emerging from the data without privileging any theorised definition. Stage 3 aimed to validate the factor structure found in Stage 2 in an independent student sample using confirmatory factor analysis (CFA). This study received ethical approval from the University of Sussex Sciences and Technology Cross-Schools Research Ethics Committee.

Stage 1: Item Generation and Review

Method

Participants. Three groups of experts were consulted to review and contribute to the items from existing self-report measures of compassion: 1) experts in teaching

contemplative approaches to others (mindfulness or compassion-based interventions), 2) experts in delivering care to others in healthcare or other pastoral care settings (e.g., healthcare staff, teachers, university lecturers), and 3) experts by experience (e.g., recipients of healthcare). Experts were recruited through e-mail invitations to an established mental health service user research group, the UK Network for Mindfulness-based Teacher Training Organisations, and faculty at the School of Psychology at the host University. Fifteen experts (66.67% female, 86.67% Caucasian) completed the consultation: eight experts in teaching contemplative approaches to others, five experts in delivering care to others in healthcare or other pastoral care settings, and two recipients of healthcare. Experts' ages ranged from 35 to 77 years (M = 52.80, SD = 11.52).

Compassion measures. Although nine compassion measures were identified in Strauss et al.'s (2016) review, only items from the following four measures were included in this consultation: the compassionate love scale (CLS) (Sprecher & Fehr, 2005), Pommier's (2010) compassion scale (CS-P), relational compassion scale (RCS) (Hacker, 2008), and Martins, Nicholas, Shaheen, Jones, and Norris' (2013) compassion scale (CS-M). Excluded from the consultation were the Santa Clara brief compassion scale (SCBCS) (Hwang, Plante, & Lackey, 2008), compassionate care assessment tool (CCAT) (Burnell & Agan, 2013), Schwartz Center compassionate care scale (SCCCS) (Lown, Muncer, & Chadwick, 2015), self-compassion scale (SCS) (Neff, 2003b), and short form of the self-compassion scale (SCS-SF) (Raes, Pommier, Neff, & Van Gucht, 2011).

The SCBCS was not included because it is a shortened version of the CLS. Items from the CCAT and the SCCCS were not included because they measure patients' ratings of compassionate care received from healthcare staff and were not self-report measures of compassion. The SCS and SCS-SF were not included for several reasons. First, although compassion for others and self-compassion have been theorised to be part of the same overarching construct, research findings have generally not supported a relationship between these two constructs. Second, items from the SCS and SCS-SF could not be easily reworded to apply to compassion for others, or compassion more generally. Lastly, the CS-P, which was included in this study, was developed based on the factor structure of the SCS and included the same elements (kindness, mindfulness, common humanity). Therefore, although items in the CS-P and

SCS are not the same, the way in which these measures conceptualise the key elements underpinning compassion for others and self-compassion is the same.

Compassionate love scale (Sprecher & Fehr, 2005). The 21-item CLS measures compassionate or altruistic love for close others and all of humankind, including strangers. Strauss et al. (2016) noted that the CLS includes items related to four of the five elements in their definition: emotionally connecting with other people's suffering, understanding something about their experience/suffering as a fellow human being, accepting and not judging them (implying tolerance), and being motivated to help them.

Compassion scale (Pommier, 2010). The 24-item CS-P mirrors the factor structure of Neff's (2003b) SCS and consists of the same three subscales: kindness versus indifference, common humanity versus separation, and mindfulness versus disengagement. The CS-P contain items which capture four of the five elements in Strauss et al.'s (2016) definition: recognising suffering, emotionally connecting with another person's distress, understanding their experience as a fellow human being, and being motivated to act or acting to alleviate suffering.

Compassion scale (Martins et al., 2013). The 10-item CS-M was developed to tap into five aspects of compassion: generosity, hospitality, objectivity, sensitivity, and tolerance across social networks and relationships. Two versions of each item exist. For the first three questions, one version relates specifically to friends and the other to strangers. For the fourth question, one version relates to friends and the other to family, and for the fifth question, one version relates to the self and the other version to other people. Therefore, this scale consists of five unique items. CS-M items focus exclusively on compassionate acts and therefore capture only the acting to alleviate suffering element of Strauss et al.'s (2016) five-element definition.

Relational compassion scale (Hacker, 2008). The 16-item RCS consists of four subscales measuring respondents' compassion for others, self-compassion, beliefs about how compassionate other people are to each other, and beliefs about how compassionate other people are to them. The four items from the beliefs about how compassionate other people are to each other subscale were not included in the consultation because they do not directly involve the participants themselves. RCS items capture four of the five elements from Strauss et al.'s (2016) review: recognising suffering, accepting and not judging others (implying tolerance), emotionally connecting with their distress, and acting to alleviate suffering.

Procedure. In order for experts to review the content validity of each item from existing compassion measures and generate additional items (using Strauss et al.'s (2016) definition as a guide), prior to consultation, two researchers in contemplative approaches met with the first author and designated each item from the included self-report compassion measures to one of the five elements from Strauss et al.'s definition. Disagreements were resolved through discussion and there was 100% agreement for the final allocations.

Where strictly necessary, the two experts and the first author met together to reword items to overcome the limitations highlighted by Strauss et al. (2016) and to fit the format of the question and response scale chosen by the research team ("Thinking about yourself in general, indicate how true the following statements are of you by choosing the appropriate number on a scale from 1 (not at all true of me) to 7 (completely true of me)"). Examples include changing items worded as questions to statements, altering items to refer to all other people (not specifically close others, friends, or strangers), and removing frequency terms (e.g., sometimes, often). However, changes to wording were kept to an absolute minimum; we endeavoured to keep as many items as possible in their original form. A total of 62 items were gathered from existing compassion measures.

The online consultation, implemented on Bristol Online Surveys (BOS; www.onlinesurveys.ac.uk), provided experts with Strauss et al.'s (2016) five-element definition and instructed them to: A) decide for each element whether existing items adequately represented this aspect of compassion, by selecting 'yes' or 'no', and B) suggest up to five additional items for each element if they thought that any were missing, using open text boxes. Experts could also leave general comments about the items designated to each element. It was decided *a priori* that in part A), for each item, at least 50% of experts needed to respond 'yes' to demonstrate that the item adequately represented the element; if more than half of experts agreed that an item was not a good indicator of an element, that item would be removed for Stage 2. It was also decided *a priori* that in part B), suggested items would be included if they: 1) did not semantically overlap with an existing item (i.e., where the wording was different but the meaning the same or very similar) and 2) adequately captured the relevant element of compassion. Two members of the research team reviewed each suggested additional item and came to an agreement concerning whether or not items should be included.

Results

Based on the feedback from experts in part A), eight of the 62 items were removed because they were not deemed to adequately represent aspects of compassion by over 50% of respondents. These were: “I like to listen to other peoples' experiences’” (RCS), “I feel a selfless caring for other people” (CLS), “I feel considerable compassionate love for people around me” (CLS), “I would rather suffer myself than see someone else suffer” (CLS), “I would rather engage in actions that help others than engage in actions that would help me” (CLS), “I would be willing to do the right thing even if it puts others at risk” (CS-M), “I would be willing to allow others pleasure of something even if it caused me pain” (CS-M), and “I don’t think much…” (CS-P; for the full item, refer to Pommier (2010)). Twenty-six items were added based on experts’ suggestions in part B). Therefore, in total, the pool of compassion items for Stage 2 of this study comprised of 80 items (refer to Appendix E for the pool of compassion items).

Stage 2: Exploratory Factor Analysis

Method

Participants and procedure. Participants were 206 University students (77.18% female). Their ages ranged from 18 to 50 years (M = 22.30, SD = 4.67). Inclusion criteria were that participants must either be undergraduate or postgraduate students at the host University. There were no exclusion criteria. Participants completed a survey containing the 80 compassion items derived from Stage 1 of this study in exchange for course credits or entry into a prize draw. The survey was hosted on BOS and items were arranged such that they alternated among the five elements.

The online survey was part of a larger study and also contained the following self-report measures: the 24-item Five Facet Mindfulness Questionnaire (FFMQ) (Bohlmeijer, Peter, Fledderus, Veehof, & Baer, 2011), 12-item Self-Compassion Scale (Raes et al., 2011), 21-item Depression, Anxiety, and Stress Scale (Henry & Crawford, 2005), the 7-item Short Warwick-Edinburgh Mental Wellbeing Scale (Stewart-Brown et al., 2009), and the 21-item Interpersonal Reactivity Index (Davis, 1980).

Sample size. Rules of thumb relating to sample size for EFA are generally regarded as not useful or valid (MacCallum, Widaman, Zhang, & Hong, 1999). Instead,

studies have shown that adequate sample size should be determined by the nature of the data. MacCallum et al. stress the importance of level of item communality, or the proportion of variance in a variable shared with other variables (Field, 2013), in determining sample size. With a sample size of 206 participants, the factor structure should be stable provided communalities are around .50 or greater and factors are well-determined (at least three items per factor and strong loadings of items to factors) (MacCallum et al., 1999).

Statistical analyses. All analyses were conducted using SPSS version 22 (IBM, 2013). Negatively-phrased items were reverse-coded prior to analysis. Preliminary analyses involved examining the intercorrelation between compassion items to identify and remove variables which did not correlate with any other variables or correlated highly with other variables ($r > .90$). Additionally, Bartlett's test of sphericity was checked, and sampling adequacy was examined using the Kaiser-Meyer-Olkin (KMO) measure to assess the suitability of the data for EFA. EFA was conducted using principal components analysis (PCA), and oblique rotation (direct oblimin) was selected to allow for correlations among the factors.

Within each identified factor, the highest loading items (with factor loadings of at least .50) were selected to be tested in the CFA in Stage 3. A minimum of three and maximum of five items were selected per factor to ensure that factors were well-determined and of manageable length. Items which were not theoretically related to the highest loading items in a factor were not included.

Results

The initial PCA yielded 16 factors with eigenvalues greater than 1.0, which collectively accounted for 68.98% of the total variance. However, the scree plot indicated that a five-factor solution should be extracted. Costello and Osborne (2005) recommend using the scree plot to decide the number of factors to retain, because including all factors with eigenvalues greater than 1.0 is one of the least accurate methods (Velicer & Jackson, 1990). Therefore, a second PCA was conducted which specified the extraction of five factors.

The five-factor solution explained 49.96% of the total variance. There were no variables with problematically low correlations or correlated highly with other variables ($r > .90$). The overall KMO value was .92, KMO values for individual

variables were > .50, and Bartlett's test of sphericity was significant ($p < .001$). These results indicate that EFA was appropriate for the data. The factor structure and loadings of items to factors are presented in Table 16. The origin of each item is also given (CLS, CS-P, RCS, CS-M, or expert consultation). Only items with strong factor loadings (.50 or greater) and low loadings on all other factors (a difference of at least .20 between the highest loading and loadings onto other factors) are included in Table 16. Most item communalities were greater than .50, which indicates that the sample size was adequate (MacCallum et al., 1999).

Items in factor 1 generally appear to represent motivation to act or acting to help/alleviate suffering, factor 2 represents tolerating uncomfortable feelings so that we are able to help, factor 3 represents understanding the universality of suffering in human experience, factor 4 represents recognising suffering, and factor 5 represents emotionally connecting with the person in distress. These factors support Strauss et al.'s (2016) five-element definition. Asterisks next to items in Table 16 indicate selected items for CFA.

Table 16.

Factor structure and loadings of compassion items in a sample of 206 students

Item source and content	Factor loading[a]				
	1	2	3	4	5
Factor 1: Acting to help/alleviate suffering					
*CLS: If given the opportunity, I am willing to make sacrifices in order to let other people achieve their goals in life.	.73				
*CLS: If a person needs help, I would do almost anything I could to help him or her.	.67				
*CLS: I want to spend time with others so that I can find ways to help enrich their lives.	.65				
*CLS: One of the activities that provides me with the most meaning to my life is helping others.	.65				
*EC: If someone is suffering I go out of my way to help them if I can.	.62				
CLS: When someone is troubled, I feel extreme tenderness and caring.	.59				
CLS: I spend a lot of time concerned about the well-being of other people.	.58				
CLS: When I see people feeling sad, I feel a need to reach out to them.	.57				
RCS: When other people are emotionally upset I treat them with kindness and care.	.52				
Factor 2: Tolerating uncomfortable feelings					
*EC: When I see someone feeling upset I feel so overwhelmed by my emotions that I find it difficult to help them. [R]		.64			
*EC: When someone is suffering it can be hard to help them because it is so upsetting. [R]		.59			
*EC: I get carried away by my own emotional response to other people's problems or suffering. [R]		.58			
Factor 3: Understanding the universality of suffering					
*CS-P: I believe that suffering is just… [b]			.77		
*CS-P: Despite my differences with others, I know… [b]			.68		

*CS-P: It is important to me to recognize that… [b]	.60		
CS-P: When people tell me about their problems, I… [b]	.56		
*CS-P: I know that everyone feels down… [b]	.52		
Factor 4: Recognising suffering			
*CS-P: I notice when people are upset… [b]		-.75	
*EC: I notice when someone is different from how they usually are.		-.74	
*EC: I find it easy to recognise when someone is suffering or in need.		-.67	
*EC: I find it difficult to notice when people are upset. [R]		-.67	
*EC: I can understand how people are feeling even if I do not identify with their experiences.		-.55	
Factor 5: Emotional connection			
*CS-P: When people talk about their problems… [b R]			-.76
*CS-P: I feel detached from others… [b R]			-.71
CS-P: When others are feeling troubled, I… [b R]			-.70
*CS-P: I don't feel emotionally… [b R]			-.67
*EC: It is hard for me to relate to others when I see them suffering. [R]			-.67
CS-P: I tune out when people… [b R]			-.66
*CS-P: I don't concern myself… [b R]			-.65
CS-P: I try to avoid people who… [b R]			-.62
CS-P: I am cold to… [b R]			-.59
EC: If someone is in distress or trouble, I wait for other people to respond first. [R]			-.57
CS-P: I can't really connect with… [b R]			-.57
CS-P: When I see someone feeling down, I… [b R]			-.56
CS-P: When people cry in front of… [b R]			-.55

CLS = Compassionate Love Scale; CS-P = Pommier's Compassion Scale; EC = expert consultation; RCS = Relational Compassion Scale.

[a] Items with factor loadings of less than .50 or cross-loadings (a difference of less than .20 between the highest loading and loadings on other factors) are suppressed.

[b] Only item stems (50% or less of the full items) are given for items from the CS-P. For full items, please refer to Pommier (2010).

[R] Items are negatively-phrased and have been reverse-coded prior to analysis.

*Item included in the confirmatory factor analysis in Stage 3.

Stage 3: Confirmatory Factor Analysis

Method

Participants and procedure. Participants were a new sample of 256 undergraduate and postgraduate students at the host University (81.64% female, 84.38% Caucasian). Their ages ranged from 18 to 50 years (M = 19.87, SD = 4.08). Participants completed a survey on BOS containing the selected compassion items from Stage 2 of this study in exchange for course credits. The online survey was part of a larger study and also contained the Big Five Inventory measure (John & Srivastava, 1999).

Statistical analyses. Three CFA models were tested using maximum-likelihood estimation with robust standard errors (MLR) conducted in Mplus version 6. MLR was used because it produces standard errors and a chi-square test statistic that are robust to deviations from normality. First, a one-factor model in which all item are indicators of a single compassion factor was tested. Next, the five-factor model derived from the EFA in Stage 2 was tested. Lastly, a hierarchical model was tested, in which the five factors are components of a broad compassion factor. Six fit indices were used to evaluate the fit of the models to the data: the comparative fit index (CFI) (Bentler, 1990), the root mean square error of approximation (RMSEA) (Steiger, 1990), the non-normed fit index (NNFI) (Bentler & Bonett, 1980), the standardised root mean square residual (SRMR), the Akaike information criterion (AIC) (Akaike, 1974), and the relative chi-square test.

As a rule of thumb, CFI and NNFI should be greater than or close to .95 to indicate acceptable fit (Hu & Bentler, 1999), an RMSEA value of .05 or less is considered a good fit, .08 indicates acceptable fit, and .10 or more a poor fit (Browne & Cudeck, 1993), and the SRMR should be less than .08 for acceptable fit (Hu & Bentler, 1999). Relative chi-square values were obtained by dividing the chi-square test statistic by the degrees of freedom. Good fit is indicated by relative chi-square values of less than or equal to 2 and acceptable fit is indicated by values between 2 and 3 (Schermelleh-Engel, Moosbrugger, & Müller, 2003). The significance of the chi-square statistic was not used as a primary measure of model fit because of its hypersensitivity (e.g., to large sample sizes, non-normality, and large correlations between variables) (Kline, 2011). The AIC was used to compare the fit of the three models, with lower

values indicating superior fit. In addition to these fit indices, the significance of factor loadings was also considered when interpreting which model provided a better fit to the data.

Results

Table 17 shows the fit indices for the three CFA models. Bold indices (CFI, RMSEA, NNFI, SRMR, and relative χ^2) indicate acceptable fit. All indices show that the fit of the one-factor model to the data was poor, which suggests that items are not direct indicators of an overall compassion factor. RMSEA, SRMR, and relative χ^2 values for the five-factor model indicated good model-data fit. However, CFI and NNFI values for this model were just under the threshold for acceptable fit. Similarly, for the five-factor hierarchical model, RMSEA, SRMR, and relative χ^2 values indicated good fit but CFI and NNFI values suggested marginally-acceptable fit. In the five-factor hierarchical model, loadings of all five factors to the overarching compassion factor were significant (Table 18). This suggests that the five factors can be seen as elements of an overall compassion construct. All loadings of items on to relevant factors in both the five-factor and five-factor hierarchical model were also significant. Based on both the fit indices and significance of factor loadings, the five-factor hierarchical model can be interpreted as providing a better fit compared to the other two models.

Table 19 displays the factor intercorrelations in a five-factor model. All intercorrelations were significant with the exception of the correlation between the 'tolerating uncomfortable feelings' and 'acting to alleviate suffering' factors.

Table 17.

Fit indices for the three CFA models tested in a sample of 256 students

Model	CFI	RMSEA [90% CI]	NNFI	SRMR	Relative χ^2 (χ^2 / *df*)	χ^2 (*df*)	AIC
One-factor	.707	.095 [.087, .103]	.676	.085	3.318	693.384 (209)	16821.585
Five-factor	.937	**.045 [.035, .055]**	.927	**.057**	**1.522**	302.957 (199)	16383.586
Five-factor hierarchical[a]	.924	**.049 [.039, .059]**	.914	**.063**	**1.616**	329.716 (204)	16403.647

AIC = Akaike information criterion; CFA = confirmatory factor analysis; CFI = comparative fit index; CI = confidence interval; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual.

Bold indices (CFI, RMSEA, NNFI, SRMR, and relative χ^2) indicate acceptable fit when rounded up or down to two decimal places.

[a] Five-factor hierarchical refers to the model in which all five factors load onto an overarching compassion factor.

Table 18.

Standardised loadings of factors to an overarching compassion factor in the five-factor hierarchical CFA model (N = 256)

Factor	Standardised loading (*SE*)
Recognising suffering	0.77 (0.05)*
Understanding the universality of suffering	0.61 (0.08)*
Emotional connection	0.90 (0.04)*
Tolerating uncomfortable feelings	0.25 (0.10)*
Acting to alleviate suffering	0.82 (0.05)*

Note. *$p < .05$.

Table 19.

Factor intercorrelations in the five-factor model (N = 256)

	1	2	3	4	5
1. Recognising suffering	-				
2. Understanding the universality of suffering	.51*	-			
3. Emotional connection	.68*	.51*	-		
4. Tolerating uncomfortable feelings	.25*	.29*	.28*	-	
5. Acting to alleviate suffering	.63*	.51*	.75*	.01	-

Note. *$p < .05$.

General Discussion

This current study used self-report items to empirically examine the five-element definition of compassion proposed by Strauss et al. (2016). The study consisted of three stages: a systematic consultation with expert groups to review items from existing self-report measures of compassion and generate additional items (Stage 1), EFA to identify the factor structure of compassion and remove redundant items (Stage 2), and CFA to validate the factor structure found in Stage 2 (Stage 3).

Findings from the EFA supported a five-factor structure of compassion consistent with Strauss et al.'s (2016) five-element definition. The five factors emerging from the analysis were: 1) recognising suffering, 2) understanding the universality of suffering in human experience, 3) emotionally connecting with the person in distress, 4) tolerating uncomfortable feelings so that we are able to help, and 5) being motivated to act or acting to help/alleviate suffering. Findings from the CFA provided promising support for this five-factor model and a five-factor hierarchical model, with all but two indices demonstrating good fit and significant item loadings and loadings of all five factors to an overarching compassion factor.

However, inspection of factor intercorrelations (in a five-factor model) and the magnitude of loadings of factors to an overarching compassion factor (in a five-factor hierarchical model) highlight issues with the 'tolerating uncomfortable feelings' factor, which did not correlate significantly with the 'acting to alleviate suffering' factor and had the smallest factor loading (standardised loading = .25; standardised loadings of other factors = .61 to .90). The non-significant correlation between tolerating and acting factors was unexpected, given the inclusion of both elements in theoretical conceptualisations of compassion (e.g., Feldman & Kuyken, 2011; Gilbert, 2010), and it is uncertain why tolerating would be significantly related to all factors but the acting factor. These findings could suggest that tolerating uncomfortable feelings is not a core part of the compassion construct. However, it would be premature to exclude the 'tolerating' factor from the five-element definition at this stage, given its inclusion in conceptualisations of compassion, significant factor loading, and significant intercorrelations with four of the five factors. We would therefore recommend that further empirical research explores the relationship between tolerating uncomfortable feelings and compassion more broadly and with its proposed subcomponents, the outcome of which may be that 'tolerating' is excluded from future definitions.

In addition to this factor-level issue, there were a number of issues with the items from existing measures loading onto the identified factors. One limitation is that not all items appear to be measuring compassion; for example, some items do not refer specifically to suffering (e.g., "I want to spend time with others so that I can find ways to help enrich their lives") and others appeared to be measuring self-sacrificing rather than mutual compassion for self and others (e.g., "If given the opportunity, I am willing to make sacrifices in order to let other people achieve their goals in life"). Such items do not appear to be describing compassion, but instead related constructs such as kindness and altruism (see Strauss et al. (2016) for a discussion of the distinction between these constructs).

Another limitation is that item wording is inconsistent across different measures (reflecting that they were derived from a range of existing measures). Additionally, only three items, which were generated through expert consultation, loaded strongly on the 'tolerating' factor. Similarly, only four theoretically related items loaded strongly onto the 'universality of suffering' factor. This can be attributed to the paucity of items in existing self-report measures which capture these dimensions of compassion.

Limitations at the item-level such as these lead us to recommend against using these items collectively to measure compassion. A measure which includes these items is unlikely to be fully and coherently representative of compassion. Instead, we recommend the development of a new self-report questionnaire, following good practice guidelines for measure development, that is rigorously tested for its psychometric properties. Limitations at the item-level also preclude strong conclusions at this early stage regarding the conceptual structure of compassion and whether or not this should include the 'tolerating uncomfortable feelings' factor. Although current findings suggest that the 'tolerating' factor may be problematic, given its small factor loading and the lack of correlation with the 'acting' factor, we recommend that future research developing a new measure of compassion, which generates new items in order to overcome the item-level limitations noted above, include tolerating items in their initial item pool alongside items representing the other four elements from Strauss et al.'s (2016) definition. This inclusive approach will allow future research to examine the factor structure emerging from the data, and whether tolerating is a part of this, in the absence of item-level limitations. If future research replicates current findings and highlights 'tolerating' as a problematic factor, then this would indicate that compassion may be better represented by four rather than five elements.

Strengths and Limitations

Strauss et al. (2016) reviewed the theoretical literature on compassion and consolidated the range of conceptualisations into one multifaceted definition of compassion. Although factor-level and item-level limitations have been identified, present findings provide preliminary empirical support for a five-factor hierarchical model of compassion consistent with Strauss et al.'s definition. These findings contribute to a greater understanding of the construct and provide an empirically-supported foundation for future measure development. However, the following limitations should be taken into account.

Strauss et al. (2016) stated that their definition of compassion could be applied to both the self and other people. However, the current study excluded measures of self-compassion (SCS and SCS-SF) for the following reasons, also stated in the Stage 1 method section: items from the SCS (Neff, 2003b) and SCS-SF (Raes et al., 2011) could not be easily reworded to apply to compassion towards others or compassion more generally (without specifying a target), Pommier's (2010) compassion scale and the SCS share the same factor structure (so the key elements in the SCS are represented in Pommier's scale), and there is a lack of empirical evidence supporting a relationship between self- and other-compassion. Therefore, current findings support Strauss et al.'s definition of compassion directed towards other people, but not towards the self. To advance our understanding of the conceptual structure of compassion and the relationship between self- and other-compassion, future efforts to develop a new measure of compassion based on Strauss et al.'s definition should attempt to generate parallel items which can be applied to the self or others. Factor analyses could then illuminate whether the factor structure of compassion is the same when it is applied to others as it is when it is applied to the self. In addition, this approach would help to clarify the empirical relationship between self-compassion and other-compassion, as parallel items should maximise the possibility of demonstrating a relationship between the two if one indeed exists. This line of research therefore has the potential to yield insights into the nature of the relationship and overlap between self-compassion and other-compassion.

The items in the 'emotional connection' and 'tolerating uncomfortable feelings' factors were all negatively-phrased and reverse-coded prior to EFA. Although such

items were negatively-phrased, we labelled their factors positively (i.e., 'emotional connection' rather than 'emotional disconnection') so that the labels for all factors would be in the same direction and consistent in indicating compassion, rather than lack of compassion (cf. the labelling of judging items as 'non-judging' in the FFMQ (Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006)). By doing this, we are not asserting that, for example, emotional connection is the exact converse of emotional disconnection; this would need testing empirically. Our aim was simply to examine the factors emerging from exploratory factor analysis of existing compassion items. This is an initial step in the process of empirically validating the conceptual structure of compassion and providing a foundation for future measure development work. Limitations of the items which emerged from the analysis, including a reliance of negatively-phrased items for certain factors, strengthens our proposal for a new measure of compassion with newly developed items. We recommend that future research developing a new measure of compassion generate primarily positively-worded items, to reduce the possibility of factors emerging which consist solely of negatively-worded items and may not be ideally captured by a positively-phrased label.

EFA and CFA of compassion items were also limited to data from samples of University students. Using different samples (e.g., meditators, older populations, healthcare professionals) could result in greater support for the five-factor structure of compassion or the identification of alternative factor structures. However, even if further research replicated the current study using samples from more diverse populations, it is likely that similar item-level limitations will be present (e.g., poor content validity of items, inconsistent wording of items from different measures, fewer items loading onto particular factors). Therefore, we recommend that future research develop a new measure of compassion, in order to overcome the item limitations highlighted here. The factor structure of this new measure should then be tested widely in different populations.

Future Research

Findings from this study open up new avenues of research into compassion, which we are currently exploring. The primary focus of future research should be to develop a psychometrically strong self-report measure of compassion using Strauss et al.'s (2016) comprehensive five-element definition to guide item generation. The

development of a high-quality measure which captures the key dimensions of compassion will bring us closer to being able to answer the important question of whether compassion can be cultivated through compassion-based interventions. In order to answer this important question, we first need a valid and reliable measure of compassion for the self and others in order to ascertain the effect of compassion-based interventions on each of these aspects of compassion. Such a measure may also help to optimise the effectiveness of these interventions by identifying the elements requiring greater therapeutic attention. Moreover, if this line of research supports the Buddhist perspective and demonstrates that enhancing self-compassion supports the development of compassion towards others, then compassion-based interventions may be more effective in cultivating compassion towards others if they begin by enhancing self-compassion. Conversely, if emerging research does not support a relationship between compassion for others and self-compassion, then we cannot assume that interventions designed to cultivate one form of compassion would necessarily enhance the other form and interventions would need to clarify their focus (in terms of whether they are designed to cultivate self-compassion, compassion for others, or both).

Although the current study supported Strauss et al.'s (2016) five-element definition, a definitive answer to the question of definition is still emerging, and theoretical and empirical developments (including the development of a new scale which comprehensively captures the elements of compassion) may also shed more light on which elements are integral to compassion (and whether 'tolerating uncomfortable feelings' is an integral element of compassion) and how elements interact to give rise to compassion.

Future research could also explore non-self-report methods of assessing compassion (e.g., behavioural or physiological measures) and triangulate these alternative measures with a new self-report measure of compassion.

Conclusions

The scientific study of compassion is in its infancy and many key questions remain poorly answered, including the fundamental question of measurement. As an essential first step to advancing our understanding of compassion and furthering research in this field, the current study empirically examined the conceptual structure of compassion using self-report items. Findings showed preliminary support for a five-

factor structure of compassion consistent with Strauss et al.'s (2016) definition; compassion consists of recognising suffering, understanding the universality of suffering in human experience, emotionally connecting with the person who is suffering, tolerating uncomfortable feelings aroused, and being motivated to act or acting to alleviate suffering. However, our findings indicated that the 'tolerating uncomfortable feelings' factor may be problematic and compassion may be better represented by four rather than five elements. This possibility requires further empirical testing. There were also limitations with items from included self-report measures (e.g., not all items appear to measure compassion, inconsistency in item wording across measures, lack of items representing certain elements of compassion), which lead us to recommend against using these items collectively to assess compassion. Taken together, we recommend developing a new measure of compassion, by generating self-report items for each of the five elements (including 'tolerating' items, to determine whether or not factor-level issues are resolved once item-level issues are addressed), and rigorously testing its psychometric properties in diverse populations.

Chapter 7:
Developing and Validating New Self-Report Measures of Compassion: Sussex-Oxford Compassion for Others Scale (SOCS-O) and Sussex-Oxford Compassion for the Self Scale (SOCS-S)

Chapter 7 is under peer review in *Psychological Assessment* as:

Gu, J., Baer, R., Cavanagh, K., Kuyken, W., & Strauss, C. Developing and Validating New Self-Report Measures of Compassion: Sussex-Oxford Compassion for Others Scale (SOCS-O) and Sussex-Oxford Compassion for the Self Scale (SOCS-S). *Manuscript under review.*

Author Contributions

Design:	All authors
Data Collection:	JG
Data Analysis:	JG
Initial Draft:	JG
Reviewing and Editing:	All authors

Abstract

Compassion has received increasing societal and scientific interest in recent years. The science of compassion requires a tool that can offer valid and reliable measurement of the construct to allow examination of its causes, correlates, and consequences. The current programme of research developed and validated new self-report measures of compassion for others and compassion for the self based on the theoretically and empirically supported five-element definition of compassion; compassion involves 1) recognising suffering, 2) understanding the universality of suffering, 3) feeling for the person suffering, 4) tolerating uncomfortable feelings aroused so that we remain open to and accepting of them in their suffering, and 5) motivation to act/acting to alleviate suffering (Strauss, Taylor, Gu, Kuyken, Baer, Jones, & Cavanagh, 2016). The resulting compassion measures, the 20-item Sussex-Oxford Compassion for Others Scale (SOCS-O) and 20-item Sussex-Oxford Compassion for the Self Scale (SOCS-S), were developed and validated across four stages: 1) item generation and review through expert and non-expert consultation, 2) item reduction, 3) confirmatory factor analysis (CFA) in a sample of healthcare staff to validate the factor structure of the SOCS-O and SOCS-S, and 4) CFA in a sample of University students to cross-validate their factor structure. Findings in both healthcare staff and students support the five-factor structure from Strauss et al.'s definition for both scales. Psychometric properties of both measures were good and support their use; both the SOCS-O and SOCS-S showed good internal consistency, interpretability, floor and ceiling effects, and convergent and discriminant validity.

Keywords: compassion, self-compassion, self-report, measure, questionnaire

Introduction

Compassion is considered to be an innate, evolved capacity (Darwin, 1871; de Waal, 2009; Gilbert, 2005) and has long been emphasised to be a core human virtue in major contemplative and religious traditions (Dahlsgaard, Peterson, & Seligman, 2005). Recently, there has been a surge in scientific interest in compassion and increased recognition of the importance of both compassion for others and compassion for the self across multiple sectors of society, including healthcare, education, and the justice system (e.g., American Medical Association, 2001; Compassion in Education Foundation, 2016; Department of Health, 2013; Norko, 2005). Compassion is associated with a range of adaptive and prosocial characteristics and outcomes, such as greater wellbeing (Davidson & Schuyler, 2015), happiness (Mongrain, Chin, & Shapira, 2011), and reduced depressive symptoms (López, Sanderman, Ranchor, & Schroevers, 2018), and there is growing evidence that greater compassion can be cultivated through compassion-based interventions (Kirby, Tellegen, & Steindl, 2017). The science of compassion requires a tool that can offer valid and reliable measurement of the construct to allow examination of its causes, correlates, and consequences (Strauss, Taylor, Gu, Kuyken, Baer, Jones, & Cavanagh, 2016). This paper reports on the development and psychometric properties of parallel measures of compassion for others and compassion for the self.

While there are many definitions of compassion, there has been a lack of consensus on its key defining features. In a recent position paper, we reviewed and consolidated a range of theoretical conceptualisations of compassion into one multifaceted, operational definition in an attempt to provide the clarity necessary to advance compassion research (Strauss et al., 2016). Our definition conceptualises compassion as a cognitive, affective, and behavioural process consisting of the following five elements: 1) recognising suffering, 2) understanding the universality of suffering in human experience, 3) feeling for the person suffering and emotionally connecting with their distress, 4) tolerating any uncomfortable feelings aroused in response to the suffering (e.g., fear, disgust, distress) so that we remain accepting of and open to the person suffering, and 5) acting or being motivated to act to alleviate the suffering. As well as encompassing these elements, a key feature of compassion that distinguishes it from related states (e.g., empathy, kindness, sympathy) is that it arises specifically in

response to suffering (Strauss et al., 2016). Consistent with theory that the process of compassion is broadly the same whether it is directed at the self or at others (Feldman & Kuyken, 2011; Gilbert, 2009b, 2014), this five-element definition applies to both. That is to say recognising suffering, and its universality, being able to tolerate elicited feelings, and acting to alleviate suffering can be directed equally to the self or others.

This five-element definition of compassion is supported both theoretically and empirically. The definition is consistent with various contemporary theoretical accounts of compassion's role in emotion regulation, interpersonal relating, and prosocial behaviour (e.g., Gilbert, 2005; Goetz, Keltner, & Simon-Thomas, 2010; Singer & Klimecki, 2014). The five elements have also received empirical support through exploratory and confirmatory factor analyses of items from existing self-report compassion scales (Gu, Cavanagh, Baer, & Strauss, 2017). Although factor analyses of existing items support the five-element definition of compassion, existing measures of compassion fail to capture the breadth of all five elements, are limited in their reliability and/or validity, have limitations with item wording or response scales, and have been designed to capture either self- or other-compassion, rather than both (see Strauss et al. 2016 for a detailed review).

Given the lack of valid and reliable measures which comprehensively capture compassion, there is a need to develop new robust measures of compassion for the self and others in order to progress scientific investigation. Continued use of measures which are limited both in whether they fully capture the nature of compassion and in their psychometric properties could lead to erroneous research findings which would be counterproductive for compassion research and practice. Key areas of research which would benefit from new robust measures of compassion include evaluating the causes and consequences of compassion and examining whether psychological interventions developed to explicitly or implicitly enhance people's capacity for compassion for themselves and other people (e.g., mindfulness-based interventions, Compassion Focused Therapy; Gilbert, 2014, Mindful Self-Compassion; Neff & Germer, 2013) work through their hypothesised mechanism of action (i.e., improved compassion).

The Current Programme of Research

The current programme of research aimed to address the lack of robust measures by developing and psychometrically evaluating two parallel self-report measures of compassion based on Strauss et al.'s (2016) theoretically and empirically supported five-element definition of compassion; the Sussex-Oxford Compassion for Others Scale (SOCS-O) and the Sussex-Oxford Compassion for the Self Scale (SOCS-S). Self and other versions of the scale were developed in parallel in keeping with the theoretical literature on compassion which does not distinguish between the two (e.g., Feldman & Kuyken, 2011; Gilbert, 2009b, 2014; Strauss et al., 2016). Developing compassion for self and other scales in parallel has the potential to empirically test this theory and to enhance understanding of the nature of the relationship between compassion for the self and compassion for others (Gu et al., 2017). Parallel scales will clarify the facets underlying compassion for self and others (theory would predict that the factor structure of both scales will mirror each other) and will also enable empirical examination of the overlap between the experience of compassion for self and others.

Development and validation of the SOCS-O and SOCS-S comprised four stages: 1) item generation and review through consultation with both experts and non-experts, 2) item reduction using data from a sample of healthcare staff, 3) validation of the factor structure of measures and evaluation of their psychometric properties in a sample of healthcare staff, and 4) cross-validation of their factor structure and evaluation of their psychometric properties in a sample of University students. Healthcare staff were recruited in Stages 2 and 3 for a number of reasons. First, they represent a well-defined sample for whom compassion for self and others may be particularly salient on a daily basis, given their experience of providing care to others whilst working in an emotionally demanding profession. Second, in response to increasing research and societal interest in compassion in healthcare contexts; there has been a particular emphasis on creating a culture of compassion in the healthcare sector (e.g., American Medical Association, 2001; Department of Health, 2013; NHS England, 2017), which is linked to research indicating improvements in patient outcomes associated with increased compassionate care (e.g., Epstein et al., 2005; Sanghavi, 2006), acknowledgement of self-care as an integral part of providing effective care to others (NHS England, 2017), and reports of diminishing compassion for self and others in cases of work-related burnout (Joinson, 1992). Lastly, recruiting

healthcare staff allowed for empirical testing of key research questions in this sample, including whether compassion for the self is related to providing compassionate care to others and whether enhanced compassion is linked to reduced work-related burnout. These questions were addressed in Stage 3 of this programme of research.

The four stages followed best practice guidelines for measure development in terms of generating items in relation to a theoretically informed, operational definition and in consultation both with experts in the topic and non-experts from a population likely to complete the measures in future research, assessing the content validity of items, reducing item pools to remove redundant items and create scales of manageable length, validating factor structures in independent samples to confirm a prespecified model for the measures, and assessing other psychometric properties, such as internal consistency and convergent and discriminant validity (e.g., Byrne, 2005; Costello & Osborne, 2005; Furr, 2011; Hinkin, 1998). The method and results for each stage are presented in turn.

Stage 1: Item Generation and Review

To maximise content validity, we used the five-element definition of compassion to formulate items that closely related to each dimension. Items relating to self-compassion and other-compassion were generated in parallel. Items were generated and revised in consultation with experts in contemplative approaches purposively sampled to represent different cultural contexts across the globe and reviewed by non-experts representative of the populations likely to complete the measure.

Item Generation

Method.

Participants and procedure. 22 English speaking experts in contemplative approaches (72.7% female; M_{age} = 43.50 years, SD_{age} = 11.62), defined as researchers and/or teachers in the fields of mindfulness or compassion with personal experience of contemplative practice (i.e., experience of cultivating mindfulness and/or compassion through contemplative meditation practices), were consulted to generate compassion items under the five elements identified by Strauss et al. (2016) and Gu et al. (2017). Experts responded to e-mail invitations distributed through contemplative research and

teacher networks. Experts had on average 10.86 years of personal contemplative practice experience (*SD* = 7.39). There were at least two experts from each of the six continents (Europe, Asia, Africa, North America, South America, Australia) and within each continental group, there was at least one representative from each expert group (researcher or teacher).

Interviews with experts were conducted by the first author over telephone or Skype. At least 24 hours prior to the interviews, experts were provided with an information sheet detailing the five elements of compassion, the interview procedure, preferred item characteristics (e.g., chosen response scale, response period, items worded as statements), and good practice guidelines for formulating items (e.g., avoiding double-barrelled items, keeping item wording concise, excluding frequency terms such as 'often' and 'sometimes'; DeVellis, 2016; Terwee, 2007). The information sheet informed experts of the intention to develop measures of both self- and other-compassion. Experts were asked to generate up to three parallel items that they thought best described each element of compassion for self and others.

Results. Altogether, experts generated 155 other-compassion items and 101 self-compassion items. All authors reviewed all generated items and came to a consensus regarding the set of items through an iterative process. In order to retain as many generated items as possible, items were removed only if they were semantic duplicates and if they did not conceptually capture a particular element of compassion. Some items were also reworded to fit the response scale and parallel items were generated where these were lacking (e.g., generating an other-compassion version of an item which had only a self-compassion form). All universality of suffering items could be applied to both the self and others (e.g., "I understand that feeling upset at times is part of human nature" and "I understand that everyone experiences suffering at some point in their lives"). Following the iterative review by authors, the pool of items was reduced to 60 compassion for others items and 60 compassion for the self items.

Item Review

Method and Results.

Fifteen of the experts in contemplative approaches who contributed to the generation of the initial pool of items and fifteen non-experts (60.0% female; M_{age} = 28.27 years, SD_{age} = 5.08) reviewed the generated items. Non-experts were

undergraduate students at a University in the South of England with no prior experience of mindfulness meditation or who have not undertaken a contemplative or compassion-based course.

An anonymous online survey containing the 60 other-compassion and 60 self-compassion items, displayed under their relevant element, was administered to participants. The survey for experts asked them to consider whether each item adequately represents its relevant element and respond accordingly by selecting 'yes' or 'no'. The survey for non-experts asked them to consider whether the wording of each item is clear and understandable ('yes' or 'no'). It was agreed a priori that an item would be removed if more than 50% of experts responded 'no', indicating that it does not adequately represent its relevant element, or if more than 50% of non-experts responded 'no', indicating that it is not clearly worded.

None of the items were removed based on the review by non-experts. More than 50% of experts responded 'no' to the following five items, indicating that they do not adequately represent their relevant elements: "I'm judgemental of others when they are going through a hard time" (other-compassion item), "When others are having difficulties, I'm cold-hearted towards them" (other-compassion item), "I overreact to other people's problems" (other-compassion item), "I misjudge how I'm feeling" (self-compassion item), and "When I'm having difficulties, I'm cold-hearted towards myself" (self-compassion item). These five items were removed, leaving 57 other-compassion items and 58 self-compassion items for Stage 2.

Stage 2: Item Reduction

Stage 2 aimed to reduce the pool of self- and other-compassion items generated in Stage 1. To do this, we applied the theoretically and empirically supported five-factor model separately on the pool of self- and other-compassion items using confirmatory factor analysis (CFA) and selected items with the highest loadings on each factor.

Method

Participants and procedure. Participants were 1,017 healthcare staff working in a role that involves at least one day a week of direct contact with patients. Staff were

recruited from public healthcare organisations in the South of the UK. Of the 1,017 participants, 859 completed demographic questions, with the exception of age, which was completed by 843 participants. The mean age of the sample was 42.37 years (*SD* = 11.99; range: 18–77 years) and 79.6% were female (*n* = 684). Most of the sample were white (90.2%) and married, in a civil partnership, cohabiting, or in a long-term relationship (73.0%). In terms of level of education, 9 (1.0%) had no formal qualifications, 80 (9.3%) had some General Certificate of Secondary Education (GCSE; UK school qualifications received at age 16) or equivalent qualifications, 145 (16.9%) had some A Levels (UK school qualifications received at age 18) or equivalent qualifications, 391 (45.5%) had a bachelor's degree or equivalent, and 234 (27.2%) had a higher degree, such as a master's or doctoral degree. The majority of staff worked in nursing (30.2%), followed by allied health (18.5%), and ambulance services (10.4%); each remaining job role category comprised less than 10% of the sample. Participants completed an anonymous online survey on Qualtrics containing several self-report measures (see below).

Measures.

Compassion items. This consisted of the 57 other-compassion items and 58 self-compassion items derived from Stage 1. The self- and other-compassion items appeared separately and their order was counterbalanced, such that for around half of participants, other- or self-compassion scales appeared first. Items were arranged such that they alternated among the five elements. Participants were instructed to indicate how true each statement was of them using a 5-point Likert scale, ranging from 1 (not at all true of me) to 5 (always true of me).

Along with the compassion items, the survey contained the following measures: the 15-item Five-Facet Mindfulness Questionnaire (Baer, Carmody, & Hunsinger, 2012), 12-item Self-Compassion Scale (Raes et al., 2011), Santa Clara Brief Compassion Scale (Hwang et al., 2008), Interpersonal Reactivity Index (Davis, 1980), 21-item Depression, Anxiety, and Stress Scale (Henry & Crawford, 2005), and Short Warwick-Edinburgh Mental Wellbeing Scale (Stewart-Brown et al., 2009). Data from these measures formed part of a wider study and are not reported on here.

Planned data analysis. Two five-factor CFA models, with items loading on respective factors from the five-element conceptualisation of compassion (Strauss et al., 2016), were applied; one to the pool of other-compassion items and one to the pool of self-compassion items. Models used maximum-likelihood estimation with robust

standard errors conducted in Mplus version 7.4 (Muthén & Muthén, 1998-2015). As the aim of this stage was to select items for the resulting scales based on their standardised loadings on factors, model-data fit indices were not reported for this stage. Examining model-data fit alongside item reduction may bias which items are selected and a stronger test would be to validate the factor structures of the resulting scales in independent samples (Stages 3 and 4) (Levine, 2005; Matsunaga, 2010). To create scales of manageable length for use in a variety of contexts, the four highest loading items were selected for each factor, creating 20-item self- and 20-item other-compassion measures.

Results

Compassion for others. 932 staff completed other-compassion items and were included in the item selection for this scale. Table 1 in Appendix F (supplementary materials) shows the standardised loadings of items on respective factors. The four highest loading items for each factor were selected for the SOCS-O and these are preceded by an asterisk. All standardised loadings were significant ($p < .001$) and all selected items had loadings greater than .40.

Compassion for the self. 947 participants completed self-compassion items and were included in the item selection for this scale. Table 2 in Appendix F (supplementary materials) presents the standardised item loadings on respective factors. The four highest loading items for each factor were retained for the SOCS-S; these are preceded by an asterisk. All standardised loadings were significant ($p < .001$) and all selected items had loadings greater than .40.

Stage 3: Validating Factor Structures using CFA

Stage 3 applied CFA to data from a large, independent sample of healthcare staff to confirm the factor structures of the SOCS-O and SOCS-S. This stage also tested other psychometric properties of these scales, namely internal consistency of total scale and subscale items (the extent to which items in a scale or subscale are correlated), convergent and discriminant validity (the degree to which scales were related to other measures in ways consistent with theoretically derived hypotheses), floor and ceiling effects (the percentage of respondents achieving the highest and lowest possible scores

on scales), and interpretability (the extent to which qualitative meaning can be attached to quantitative scores).

Method

Participants and procedure. An independent sample of 1,319 healthcare staff completed an anonymous online survey on Qualtrics containing self-report measures (see below). Staff were recruited from public healthcare organisations in the South of the UK. 1,132 to 1,137 participants completed demographic questions, with the exception of age, which was completed by 1,123 participants. The mean age of the sample was 44.83 years (*SD* = 11.30; range: 18–74 years) and 83.1% were female (*n* = 945). Most of the sample were white (89.7%) and married, in a civil partnership, cohabiting, or in a long-term relationship (76.7%). In terms of level of education, 12 (1.1%) had no formal qualifications, 144 (12.7%) had some GCSEs (UK school qualifications received at age 16) or equivalent qualifications, 201 (17.8%) had some A Levels (UK school qualifications received at age 18) or equivalent qualifications, 502 (44.3%) had a bachelor's degree or equivalent, and 273 (24.1%) had a higher degree, such as a master's or doctoral degree. The majority of staff worked in nursing (39.2%), followed by allied health services (15.2%) and administrative and clerical roles (15.3%); remaining job role categories comprised less than 10% of the sample.

Measures. With the exception of the SOCS-O and SOCS-S, the below measures were selected because they are theoretically expected to be related in particular ways to self- and/or other-compassion.

Sussex-Oxford compassion for others scale (SOCS-O) and Sussex-Oxford compassion for the self scale (SOCS-S). The 20-item SOCS-O and 20-item SOCS-S derived from Stage 2 appeared separately, either at the start or the end of the survey, and their order was counterbalanced. For each scale, items were arranged such that they alternated among the five elements. Participants were instructed to indicate how true each statement was of them using a 5-point Likert scale, ranging from 1 (not at all true of me) to 5 (always true of me). A copy of the SOCS-O and SOCS-O is included in the supplementary materials in Appendix F.

Five-facet mindfulness questionnaire 15-item version (FFMQ; Baer et al., 2012). The 15-item FFMQ (FFMQ-15) is a short form of the 39-item FFMQ (FFMQ-39) and measures the general tendency to be mindful in everyday life. It includes the

same five facets as the long form: observing, describing, acting with awareness, non-judging of inner experience, and non-reactivity to inner experience. The factor structure of the FFMQ-15 is consistent with that of the FFMQ-39, there are large correlations between total facet scores of the short and long forms, and the two FFMQ versions do not differ significantly from each other in terms of convergent validity (Gu, Strauss, Crane et al., 2016). Previous research (Baer et al., 2006; Gu, Strauss, Crane et al., 2016; M. J. Williams, Dalgleish, Karl, & Kuyken, 2014) found that in non-meditator samples, a four-factor hierarchical structure without the 'observing' facet provided a superior fit compared to a five-factor hierarchical structure. As it is likely that our current sample has little or no previous meditation experience, 'observing' items were excluded. FFMQ-15 items were rated on a 5-point Likert scale, ranging from 1 (never or very rarely true) to 5 (very often or always true). Cronbach's alpha for FFMQ-15 items (excluding observing items) was .80.

Self-compassion scale – short form (SCS-12; Raes et al., 2011). This 12-item measure is a short form of the original 26-item scale (Neff, 2003b). The SCS-12 was found to have the same factor structure as the long form, with six factors loading on a higher-order self-compassion factor: self-kindness, self-judgement, common humanity, isolation, mindfulness, and over-identification (Raes et al., 2011). Items were rated on a 5-point Likert scale ranging from 1 (almost never) to 5 (almost always), with the total score ranging from 12 to 60. Cronbach's alpha for SCS-12 items was .88.

Santa Clara brief compassion scale (SCBCS; Hwang et al., 2008). The 5-item SCBCS is a short form of the 21-item Compassionate Love Scale (Sprecher & Fehr, 2005) and measures compassion towards strangers and humankind at large. Responses to items were given on a 7-point Likert scale, ranging from 1 (not at all true of me) to 7 (very true of me). Of all the existing other-compassion measures reviewed by Strauss et al. (2016), the SCBCS was the shortest measure which obtained the highest quality rating. Cronbach's alpha for SCBCS items was .91.

Interpersonal reactivity index (IRI; Davis, 1980). The 28-item IRI is a multidimensional measure of dispositional empathy, with the following subscales: perspective taking, fantasy, empathic concern, and personal distress. Responses were given on a 5-point Likert scale, ranging from 1 (does not describe me well) to 5 (describes me very well). Following previous research (e.g., Neff & Pommier, 2013; Pommier, 2010), we excluded the fantasy subscale, because it is not regarded as

assessing a core part of empathy. Cronbach's alphas were .79 (perspective taking), .75 (empathic concern), and .76 (personal distress).

Depression, anxiety, and stress scale – short form (DASS; Henry & Crawford, 2005). The 21-item shortened version of the DASS measures the severity of core symptoms associated with depression, anxiety, and stress. Participants were asked to indicate the presence of each symptom over the past week. Responses were given on a 4-point Likert scale, ranging from 0 (never) to 3 (almost always). Cronbach's alphas were .92 (depression), .81 (anxiety), and .86 (stress).

Short Warwick-Edinburgh mental wellbeing scale (SWEMWBS; Stewart-Brown et al., 2009). The 7-item SWEMWBS is a measure of positive mental wellbeing. This measure involves rating items on a 5-point Likert scale ranging from 1 (none of the time) to 5 (all of the time). Participants were asked to rate items based on their experience over the past two weeks. Cronbach's alpha for SWEMWBS items was .89.

Maslach burnout inventory – Human services survey (Maslach, Jackson, & Leiter, 1981). The 22-item Maslach Burnout Inventory – Human Services Survey (MBI-HSS) was designed to measure work-related burnout in professionals working in the human services such as healthcare and consists of three distinct subscales: emotional exhaustion, depersonalisation, and personal accomplishment. Participants were asked about the frequency with which they have certain experiences and items were answered on a 7-point Likert scale, ranging from 0 (never) to 6 (every day). The three subscales have been found to have adequate internal consistency, test-retest reliability, and convergent and discriminant validity. The MBI-HSS was administered to a subset of participants in this sample (n = 115). Cronbach's alphas were .90 (emotional exhaustion), .75 (depersonalisation), and .78 (personal accomplishment).

Planned data analysis. Three CFA models were tested for the 20-item SOCS-O and 20-item SOCS-S: 1) a one-factor model in which all items are direct indicators of a single compassion factor, 2) a five-factor correlated model, with items loading on respective factors from the five-element definition of compassion (Strauss et al., 2016), and 3) a five-factor hierarchical model, where the five factors load on an overarching compassion factor. All CFA models used maximum-likelihood estimation with robust standard errors conducted in Mplus version 7.4 (Muthén & Muthén, 1998-2015).

The following five fit indices were used to indicate model-data fit: the comparative fit index (CFI; Bentler, 1990), root mean square error of approximation

(RMSEA; Steiger, 1990), non-normed fit index (NNFI; Bentler & Bonett, 1980), standardised root mean square residual (SRMR), and Akaike information criterion (AIC; Akaike, 1974). Rules of thumb cut-off criteria for determining adequate fit using these indices can be arbitrary and affected by numerous factors such as sample size, data distribution, and model complexity and specifications (e.g., Chen, Curran, Bollen, Kirby, & Paxton, 2008; Marsh, Hau, & Wen, 2004), such that a model may fit the data even when one or more indices suggest inadequate fit (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Consequently, researchers do not recommend their use as absolute, universally applied rules for assessing fit (e.g., Hu & Bentler, 1999; Marsh et al., 2004).

Given these considerations, following M. J. Williams et al. (2014), we used both liberal and conservative cut-off points for acceptable fit for the CFI, RMSEA, NNFI, and SRMR: the CFI and NNFI should be close to or greater than .90 (liberal) or .95 (conservative), RMSEA should be .10 or less (liberal) or .06 or less (conservative), and SRMR should be less than .10 (liberal) or .05 (conservative). We also considered the significance of factor intercorrelations and loadings when interpreting model fit. The AIC was used to compare the fit of the models, with lower values indicating superior fit. Although the chi-square test of model fit was reported, the significance of this statistic was not used to indicate model fit because of its hypersensitivity (e.g., to non-normality, large sample sizes, large correlations between variables) (Kline, 2015).

Cronbach's alphas were conducted to test internal consistency of total scale and subscale items on the SOCS-O and SOCS-S; alphas between .70 and .95 indicate good internal consistency (Terwee et al., 2007), although for psychological constructs, values below .70 are acceptable (Kline, 1999). Floor and ceiling effects of the SOCS-O and SOCS-S were examined by calculating the percentage of respondents achieving the highest and lowest possible scores; less than 15% of the sample should receive the highest or lowest score (Terwee et al., 2007). Interpretability was tested by comparing SOCS-O and SOCS-S scores in at least four subgroups of participants (Terwee et al., 2007). Specifically, an independent t-test and one-way ANOVAs were conducted, and means and standard deviations reported, to examine whether total scale scores differ in relation to the following variables: gender, length of previous meditation experience, level of education, and marital status. To follow up one-way ANOVAs, Hochberg's GT2 post hoc tests were conducted, as sample sizes are likely to differ across subgroups (Field, 2013).

Convergent and discriminant validity were tested by examining whether each scale correlated with the measures detailed in the measures subsection in line with predictions. For this criterion to be met, at least three quarters of results should be consistent with hypotheses (Terwee et al., 2007) and in relation to convergent validity, at least two of the correlations should be large ($r \geq .50$) (Barker et al., 2002). We predicted that the SOCS-O and SCBCS, both scales measuring compassion for others, should be significantly correlated at $r \geq .50$. Similarly, the SOCS-S and SCS-12, both measures of self-compassion, were expected to be significantly correlated at $r \geq .50$. We expected the SOCS-O to be significantly correlated with the empathic concern and perspective taking subscales of the IRI at $r \geq .50$. However, although we would expect the SOCS-O to be significantly and negatively related to the personal distress subscale of the IRI, a prediction was not made as to the size of this relationship, because unlike the other two subscales, almost all personal distress items are worded ambiguously in terms of target and can be interpreted in relation to the self rather than others (e.g. "Being in a tense emotional situation scares me" and "I sometimes feel helpless when I am in the middle of a very emotional situation"). Previous research has found just a small-moderate, negative correlation between compassion for others and the personal distress subscale of the IRI (Pommier, 2010).

Consistent with research which found significant correlations, ranging in size from small-moderate to large, between compassion for the self and mindfulness, positive mental health, and wellbeing (e.g., Durkin, Beaumont, Martin, & Carson, 2016; López et al., 2018; Neff, 2003b; Pommier, 2010), but no such relationships between compassion for others and these constructs (e.g., Durkin et al., 2016; López et al., 2018; Pommier, 2010), we predicted that there would be significant correlations between the SOCS-S and the FFMQ-15, SWEMWBS, and all subscales of the DASS at least small-moderate in size (positive for the FFMQ-15 and SWEMWBS and negative for DASS subscales). It is possible that the lack of significant correlations between compassion for others and mindfulness, wellbeing, and mental health was due to limitations of existing compassion measures (Strauss et al., 2016) and we therefore explored these findings but did not make specific predictions about the relationships between the SOCS-O and these constructs. Similarly, research has found a moderate-large, negative, significant correlation between self-compassion and burnout but no such relationship between compassion for others and burnout (e.g., Durkin et al., 2016).

We therefore expected significant, moderate-large correlations between the SOCS-S and subscales of the MBI-HSS (negative for emotional exhaustion and depersonalisation and positive for personal accomplishment) but did not make predictions for the SOCS-O and MBI-HSS.

Moreover, self- and other-compassion are theoretically overlapping constructs and the process of compassion is the same whether it is directed at the self or at others. However, research into the relationship between self- and other-compassion has found no more than a small relationship between these constructs (Durkin et al., 2016; López et al., 2018; Neff & Pommier, 2013; Pommier, 2010). It is currently unclear whether the little or no empirical overlap between self- and other-compassion is due to limitations of the measures used in these studies (e.g., Strauss et al., 2016; M. J. Williams et al., 2014) or indicates that these two forms of compassion are largely distinct. Thus, no specific hypotheses were made regarding the correlation between the SOCS-O and SOCS-S, but these findings were explored. Lastly, none of the relationships between the SOCS-O or SOCS-S and other measures were expected to correlate so highly ($r \geq .80$; Field, 2013) as to indicate that they were the same construct (e.g., compassion and empathy) or that measures were indistinguishable (e.g., SOCS-O/SOCS-S and existing compassion scales).

Results

Confirmatory factor analysis.

Compassion for others. 1,242 healthcare staff completed the SOCS-O and were included in the CFA. Table 20 shows the fit indices for the three CFA models. Bold indices (CFI, RMSEA, NNFI, SRMR) indicate acceptable fit according to liberal cut-off criteria. Almost all fit indices indicated poor fit of the one-factor model to the data, suggesting that items are not direct indicators of an overarching compassion factor. All fit indices indicated good fit of the five-factor and five-factor hierarchical models according to both liberal and conservative criteria. All loadings of items on factors in these two models were significant. All factor intercorrelations in the five-factor model were significant. In the five-factor hierarchical model, all loadings of factors on the overarching compassion factor were significant, suggesting that the five factors are elements of an overall compassion for others construct. Based on both the fit indices and significance of factor loadings, the five-factor hierarchical model can be interpreted

as best fitting the data. Table 3 in Appendix F (supplementary materials) presents the standardised loadings of items onto factors in the five-factor hierarchical model for the SOCS-O and Table 4 in Appendix F (supplementary materials) the standardised factor loadings in the five-factor hierarchical model. Table 5 in Appendix F (supplementary materials) shows the correlations between total scale and subscale scores on the SOCS-O in the healthcare staff validation sample.

Compassion for the self. 1,216 healthcare staff completed the SOCS-S and were included in the CFA. Table 20 presents the fit indices for the three CFA models. All indices suggested poor fit of the one-factor model but adequate fit of the five-factor and five-factor hierarchical models. All item loadings in the two five-factor models were significant. All factor intercorrelations in the five-factor model were significant and all factor loadings in the five-factor hierarchical model were significant, suggesting that the five factors are related and are elements of an overall compassion for the self construct. Based on both the fit indices and significance of factor loadings, the five-factor hierarchical model can be seen as best fitting the data. Table 6 in Appendix F (supplementary materials) displays the standardised item loadings in the five-factor hierarchical model for the SOCS-S and Table 4 in Appendix F (supplementary materials) the standardised factor loadings in the five-factor hierarchical model. Table 5 in Appendix F (supplementary materials) presents the correlations between total scale and subscale scores on the SOCS-S in the staff validation sample.

Internal consistency. Cronbach's alphas ranged from .74 to .94 for total scale and subscale items on the SOCS-O and from .75 to .93 for total scale and subscale items on the SOCS-S (Table 21). These values are considered adequate for measures of psychological constructs (Kline, 1999; Terwee et al., 2007).

Floor and ceiling effects. Less than 15% of the sample received the highest score (100) or lowest score (20) on the SOCS-O and SOCS-S; 0.1% and 0.2% of participants received the lowest possible score on the SOCS-O and SOCS-S, respectively, and 1.6% and 0.3% of participants received the highest possible score on the SOCS-O and SOCS-S, respectively, suggesting that both scales capture variability in responses.

Interpretability. Table 22 displays the means and standard deviations of total SOCS-O and SOCS-S scores across subgroups of participants. Females scored significantly higher on the SOCS-O compared to males, $t(1118) = 5.97$, $p < .001$, however, there was no significant difference between males and females in SOCS-S

scores, $t(1115) = 0.04$, $p = .965$. Length of previous meditation experience (four levels: no previous experience, less than a year, 1 to 5 years, over 5 years) significantly affected scores on both the SOCS-O ($F(3) = 5.53$, $p = .001$) and SOCS-S ($F(3) = 13.89$, $p < .001$). Scores on both scales were significantly lower for those without any meditation experience, compared to those with 1 to 5 years' experience and over 5 years' experience (ps < .05). Additionally, those with over 5 years' meditation experience scored significantly higher on the SOCS-S compared to both participants with less than a year's experience and those with 1 to 5 years' experience (ps < .05). In terms of level of education, there was a significant difference in SOCS-S scores only, $F(4) = 3.51$, $p = .007$. The only significant post hoc difference was between those with GCSEs (UK school qualifications received at age 16) or equivalent qualifications and those with higher degrees ($p = .023$). There was no significant effect of marital status on SOCS-O ($F(3) = 1.02$, $p = .384$) or SOCS-S scores ($F(3) = 1.22$, $p = .300$).

Convergent and discriminant validity. Table 23 shows the correlation coefficients between total scale and subscale scores on the SOCS-O and SOCS-S and other constructs. Consistent with predictions, the SOCS-O had significant and large correlations with the Santa Clara Brief Compassion Scale (SCBCS), empathic concern (IRI-empathic concern), and perspective taking (IRI-perspective taking) (rs ≥ .50), and the SOCS-S had a significant and large correlation with the Self-Compassion Scale (SCS-12) ($r \geq .50$). The SOCS-O was also significantly and negatively related to personal distress (IRI-personal distress). Additionally, the SOCS-S was significantly correlated in expected directions with mindfulness (FFMQ-15), wellbeing (SWEMWBS), and stress, anxiety, and depression (DASS), with correlations ranging from moderate-large to large in size, supporting our predictions. We did not make specific predictions for the SOCS-O and mindfulness, wellbeing, and depression, anxiety, and stress, but found significant, small-moderate correlations between the SOCS-O and mindfulness (FFMQ-15) and wellbeing (SWEMWBS), and small, but significant, negative relationships between the SOCS-O and stress and depression (DASS). As predicted, the SOCS-S was found to have significant and moderate-large correlations in expected directions with all subscales of the Maslach Burnout Inventory (MBI-HSS). Although we did not make specific predictions for the SOCS-O and MBI-HSS subscales, the SOCS-O was found to have significant, small-moderate correlations with depersonalisation (negative direction) and personal accomplishment (positive

direction). Taken together, at least three quarters of results were found to be consistent with predictions, at least two correlations were large ($r \geq .50$), and none were $r \geq .80$, providing support for the convergent and discriminant validity of the SOCS-O and SOCS-S in this healthcare staff sample.

Relationship between compassion for the self and others. Healthcare staff scored significantly higher on the SOCS-O compared to the SOCS-S, $t(1126) = 32.29$, $p < .001$, $d = 1.05$, 95% CI = 0.98 to 1.13. Table 5 in Appendix F (supplementary materials) shows the correlations between total scale and subscale scores on the SOCS-O and SOCS-S in the Stage 3 sample. Total scores were found to significantly correlate with a medium-large effect size at $r = .40$. Moreover, all SOCS-O and SOCS-S subscales were significantly correlated, with coefficients ranging between small-moderate to large in size ($r = .15$ to $.78$). However, the correlation between total scale scores may be artificially inflated given that three of the four items from the universality of suffering subscale were the same for both scales. We therefore calculated the correlation between total SOCS-O and SOCS-S scores excluding the universality subscale and found these to be significantly and moderately correlated at $r = .30$ ($p < .001$).

Table 20.

Fit indices for compassion models tested in both validation samples (Stages 3 and 4)

Scale	Sample	Model	CFI	RMSEA [90% CI]	NNFI	SRMR	χ^2 (*df*)	AIC
Compassion for others	1,242 healthcare staff (Stage 3)	One-factor	.718	.122 [.119, .126]	.685	**.089**	3338.294 (170)	42176.726
		Five-factor	**.973**	**.039 [.035, .043]**	**.968**	**.028**	466.435 (160)	38170.026
		Five-factor hierarchical[a]	**.972**	**.039 [.035, .043]**	**.968**	**.029**	475.491 (165)	38174.744
	371 students (Stage 4)	One-factor	.632	.126 [.119, .132]	.589	.107	1163.712 (170)	14200.646
		Five-factor	**.966**	**.040 [.030, .049]**	**.959**	**.045**	252.665 (160)	13104.222
		Five-factor hierarchical[a]	**.964**	**.040 [.030, .049]**	**.959**	**.047**	261.210 (165)	13103.945

Compassion for the self	1,216 healthcare staff (Stage 3)	One-factor	.638	.142 [.139, .146]	.596	.132	4360.676 (170)	51699.527
		Five-factor	**.947**	**.056 [.052, .060]**	**.937**	**.050**	775.599 (160)	46658.552
		Five-factor hierarchical[a]	**.939**	**.059 [.056, .063]**	**.930**	**.068**	871.920 (165)	46772.251
	371 students (Stage 4)	One-factor	.580	.156 [.149, .163]	.530	.155	1703.097 (170)	16986.098
		Five-factor	**.930**	**.065 [.058, .073]**	**.917**	**.069**	413.800 (160)	15362.973
		Five-factor hierarchical[a]	**.925**	**.067 [.059, .074]**	**.914**	**.084**	437.055 (165)	15380.924

AIC = Akaike information criterion; CFA = confirmatory factor analysis; CFI = comparative fit index; CI = confidence interval; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual. Bold indices (CFI, RMSEA, NNFI, SRMR) indicate acceptable fit according to liberal cut-off criteria when rounded up or down to two decimal places.

[a] Five-factor hierarchical refers to a model in which all five factors load on an overarching compassion factor.

Table 21.

Cronbach's alphas for total SOCS-O and SOCS-S items and each four-item subscale using available data from both validation samples (Stages 3 and 4)

	Compassion for others		Compassion for the self	
	1,319 healthcare staff (Stage 3)	371 students (Stage 4)	1,319 healthcare staff (Stage 3)	371 students (Stage 4)
Total scale	.94	.90	.93	.91
Recognising suffering	.89	.86	.88	.85
Understanding the universality of suffering	.92	.89	.92	.91
Feeling for the person suffering	.80	.73	.84	.84
Tolerating uncomfortable feelings	.74	.61	.75	.72
Acting or being motivated to act to alleviate suffering	.91	.86	.91	.90

Note. SOCS-O = Sussex-Oxford Compassion for Others Scale; SOCS-S = Sussex-Oxford Compassion for the Self Scale.

Table 22.

Means and standard deviations of total SOCS-O and SOCS-S scores for all participants and participant subgroups using available data from both validation samples (Stages 3 and 4)

		Total SOCS-O		Total SOCS-S	
		1,319 healthcare staff (Stage 3)	371 students (Stage 4)	1,319 healthcare staff (Stage 3)	371 students (Stage 4)
	All participants	82.16 (9.73); *n* = 1238	81.16 (8.56); *n* = 371	70.79 (11.65); *n* = 1204	69.66 (11.11); *n* = 371
Gender	Female	83.03 (9.28); *n* = 941	81.62 (8.35); *n* = 326	70.93 (11.40); *n* = 937	69.73 (11.12); *n* = 326
	Male	78.42 (10.29); *n* = 179	78.12 (9.54); *n* = 42	70.97 (12.60); *n* = 180	69.48 (11.47); *n* = 42
Length of previous meditation experience	No previous experience	81.58 (9.92); *n* = 800	80.51 (8.52); *n* = 283	70.05 (11.74); *n* = 798	68.57 (11.20); *n* = 283
	Less than a year	82.66 (8.24); *n* = 109	83.47 (8.01); *n* = 49	69.95 (10.46); *n* = 107	72.08 (11.52); *n* = 49
	1 to 5 years	84.44 (8.43); *n* = 139	82.83 (9.18); *n* = 36	73.30 (10.33); *n* = 139	74.33 (8.20); *n* = 36
	Over 5 years	84.86 (9.49); *n* = 70	85.00 (6.56); *n* = 3	78.35 (10.78); *n* = 71	77.00 (5.29); *n* = 3
Level of education	No formal qualifications	77.83 (14.17); *n* = 12	-	64.67 (17.84); *n* = 12	-
	GCSE or equivalent	80.84 (10.55); *n* = 143	-	68.40 (12.30); *n* = 141	-
	A-level or equivalent	83.26 (9.75); *n* = 200	-	71.81 (11.57); *n* = 200	-

	Degree (e.g., BA, BSc) or equivalent	82.18 (9.14); *n* = 500	-	70.91 (11.01); *n* = 500	-
	Higher degree (e.g., MA, MSc, PhD) or equivalent	82.54 (9.54); *n* = 272	-	72.04 (11.61); *n* = 271	-
Marital status	Single	81.01 (9.66); *n* = 152	81.27 (8.44); *n* = 315	69.85 (12.02); *n* = 154	69.30 (11.13); *n* = 315
	Married/civil partnership/cohabiting/long-term relationship	82.37 (9.59); *n* = 868	80.53 (9.38); *n* = 55	71.33 (11.52); *n* = 864	71.53 (10.93); *n* = 55
	Separated/divorced	82.86 (9.74); *n* = 96	-	69.71 (11.70); *n* = 95	-
	Widowed	82.23 (10.91); *n* = 14	-	69.36 (8.28); *n* = 14	-

Note. SOCS-O = Sussex-Oxford Compassion for Others Scale; SOCS-S = Sussex-Oxford Compassion for the Self Scale. Standard deviations are given in parentheses.

Table 23.

Correlation coefficients between total scores on the SOCS-O and SOCS-S and other constructs in both healthcare staff (non-italicised values) and student (values in italics) validation samples (Stages 3 and 4)

	FFMQ-15[a]	SCBCS	SCS-12	IRI-EC	IRI-PD	IRI-PT	SWEMWBS	DASS-S	DASS-A	DASS-D	MBI-HSS EE[b]	MBI-HSS D[b]	MBI-HSS PA[b]
Sussex-Oxford Compassion for Others Scale (SOCS-O)	.26*** *.19****	.65*** *.59****	.18*** *.14***	.64*** *.59****	-.16*** *-.15***	.54*** *.40****	.24*** *.25****	-.07* *-.05*	-.03 *-.05*	-.07* *-.18***	-.10	-.24*	.29**
SOCS-O: Recognising suffering	.21*** *.11**	.47*** *.36****	.15*** *.04*	.46*** *.29****	-.18*** *-.17***	.40*** *.22****	.22*** *.14***	-.04 *.09*	.01 *.09*	-.05 *-.03*	-.07	-.16	.23*
SOCS-O: Understanding the universality of suffering	.20*** *.22****	.35*** *.22****	.13*** *.20****	.38*** *.33****	-.08** *-.10**	.35*** *.23****	.19*** *.30****	-.04 *-.18****	-.09** *-.19****	-.06* *-.29****	-.03	-.28**	.16

SOCS-O: Feeling for	.19***	.67***	.12***	.69***	-.07*	.50***	.17***	-.04	-.01	-.06	-.09	-.20*	.23*
the person suffering	*.07*	*.67****	*.04*	*.65****	*.06*	*.37****	*.12**	*.06*	*.05*	*-.03*			
SOCS-O: Tolerating	.29***	.57***	.23***	.53***	-.24***	.52***	.27***	-.13***	-.08**	-.12***	-.20*	-.19*	.35***
uncomfortable	*.24****	*.38****	*.18****	*.41****	*-.27****	*.36****	*.23****	*-.18****	*-.17***	*-.22****			
feelings													
SOCS-O: Acting or	.16***	.59***	.09**	.57***	-.13***	.44***	.16***	-.01	.04	-.01	-.04	-.16	.25**
motivation to act to	*.07*	*.58****	*.04*	*.52****	*-.09*	*.35****	*.12**	*.05*	*.03*	*-.10*			
alleviate suffering													
Sussex-Oxford	.55***	.23***	.65***	.23***	-.24***	.38***	.57***	-.43***	-.36***	-.45***	-.34***	-.36***	.31**
Compassion for the	*.57****	*.11**	*.63****	*.22****	*-.29****	*.29****	*.54****	*-.39****	*-.39****	*-.49****			
Self Scale (SOCS-S)													
SOCS-S:	.31***	.22***	.30***	.23***	-.11***	.24***	.31***	-.20***	-.15***	-.18***	-.16	-.23*	.24**
Recognising	*.21****	*.07*	*.07*	*.19****	*-.04*	*.12**	*.24****	*-.08*	*-.11**	*-.19****			
suffering													
SOCS-S:	.23***	.29***	.21***	.32***	-.10**	.34***	.25***	-.12***	-.19***	-.15***	-.11	-.22*	.17
Understanding the	*.20****	*.14***	*.22****	*.34****	*-.09*	*.27****	*.26****	*-.25****	*-.17***	*-.29****			

universality of suffering													
SOCS-S: Feeling for the person suffering	.50*** *.52****	.15*** *.11**	.66*** *.63****	.14*** *.13***	-.19*** *-.26****	.31*** *.23****	.53*** *.48****	-.41*** *-.33****	-.32*** *-.34****	-.44*** *-.43****	-.27**	-.26**	.25**
SOCS-S: Tolerating uncomfortable feelings	.59*** *.63****	.16*** *.05*	.70*** *.71****	.12*** *.04*	-.35*** *-.38****	.31*** *.22****	.59*** *.53****	-.50*** *-.46****	-.41*** *-.46****	-.51*** *-.46****	-.40***	-.38***	.28**
SOCS-S: Acting or motivation to act to alleviate suffering	.50*** *.51****	.12*** *.06*	.63*** *.60****	.13*** *.12**	-.19*** *-.27****	.27*** *.20****	.54*** *.44****	-.42*** *-.30****	-.32*** *-.32****	-.45*** *-.39****	-.35***	-.32**	.28**

Note. Non-italicised values are correlations from the sample of 1,319 healthcare staff (Stage 3). Values in italics are correlations from the sample of 371 students (Stage 4).

[a] Items from the observing subscale were excluded from the total FFMQ-15 score.

[b] The MBI-HSS was administered to a subset of the Stage 3 healthcare staff sample (n = 115). Students (Stage 4) did not complete this measure.

*p < .05, **p < .01, ***p < .001.

Stage 4: Cross-Validating Factor Structures using CFA

Stage 4 applied CFA to data from a sample of University students to cross-validate the factor structures of the SOCS-O and SOCS-S. This stage also assessed each scale's internal consistency, convergent and discriminant validity, floor and ceiling effects, and interpretability.

Method

Participants and procedure. A sample of 371 undergraduate University students completed an anonymous online survey on Qualtrics containing self-report measures (see below). The mean age of the sample was 19.63 years (*SD* = 3.14; range: 18–45 years) and 87.9% were female (*n* = 326). Most of the sample were white (85.7%) and single (84.9%).

Measures. The measures used in Stage 3 were administered to students, with the exception of the MBI-HSS. Cronbach's alphas for these measures were: .80 (FFMQ-15 without observing items), .87 (SCS-12), .91 (SCBCS), .81 (IRI perspective taking), .78 (IRI empathic concern), .80 (IRI personal distress), .84 (DASS stress), .82 (DASS anxiety), .89 (DASS depression), and .86 (SWEMWBS). The SOCS-O and SOCS-S appeared separately, either at the start or end of the survey, and their order was counterbalanced.

Planned data analysis. Three CFA models were tested for the SOCS-O and SOCS-S in order to cross-validate their factor structures in this student sample: 1) a single-factor model in which all item are indicators of a single compassion factor, 2) a five-factor correlated model, with items loading on respective factors from the five-element definition of compassion (Strauss et al., 2016), and 3) a five-factor hierarchical model, where the five factors load on an overall compassion factor. All CFA models used maximum-likelihood estimation with robust standard errors conducted in Mplus version 7.4. As in Stage 3, the CFI (Bentler, 1990), RMSEA (Steiger, 1990), NNFI (Bentler & Bonett, 1980), SRMR, and AIC (Akaike, 1974) were used to indicate model-data fit and the same liberal and conservative cut-off points for interpreting fit indices were employed.

Internal consistency of total scale and subscale items on the SOCS-O and SOCS-S was tested using Cronbach's alpha. Floor and ceiling effects and

interpretability of the scales were explored using the methods outlined in Stage 3; the only difference was that data on level of education was not obtained from students to examine interpretability. Convergent and discriminant validity were tested by correlating the compassion scales with the measures outlined in the measures subsection and we used the same criteria and predictions stated in Stage 3 to assess whether this property was met.

Results

Confirmatory factor analysis.

Compassion for others. All 371 students completed the SOCS-O and were included in the CFA. Table 20 displays the fit indices for the three CFA models tested on the SOCS-O in this sample. Bold indices (CFI, RMSEA, NNFI, SRMR) indicate acceptable fit according to liberal cut-off criteria. As with the CFA findings from Stage 3, fit indices indicated poor model-fit of the one-factor model but good fit of the five-factor and five-factor hierarchical models according to both liberal and conservative criteria. All item loadings in these two models were significant. Factor intercorrelations in the five-factor model were significant and all loadings of factors on the overarching compassion factor in the five-factor hierarchical model were significant. This indicates that the five factors are related and are elements of an overall compassion for others construct. Based on both the fit indices and significance of factor loadings, the five-factor hierarchical model can be interpreted as best fitting the data. Table 3 in Appendix F (supplementary materials) shows the standardised item loadings in the five-factor hierarchical model for the SOCS-O and Table 4 in Appendix F (supplementary materials) the standardised factor loadings in the five-factor hierarchical model. Table 7 in Appendix F (supplementary materials) shows the correlations between total SOCS-O scale and subscale scores in this sample.

Compassion for the self. All 371 students completed the SOCS-S and were included in the CFA. Table 20 shows the fit indices for the three SOCS-S CFA models in this sample. As with the CFA findings from Stage 3, all indices suggested acceptable fit of the five-factor and five-factor hierarchical models, but poor fit of the one-factor model. All item loadings in the two five-factor models were significant. All factor intercorrelations in the five-factor model were significant and factor loadings in the five-factor hierarchical model were significant, indicating that the factors are related

and are elements of an overall self-compassion construct. Based on both the fit indices and significance of factor loadings, the five-factor hierarchical model can be interpreted as providing the best fit. Table 6 in Appendix F (supplementary materials) shows the standardised SOCS-S item loadings in the five-factor hierarchical model and Table 4 in Appendix F (supplementary materials) the standardised factor loadings in the five-factor hierarchical model. Table 7 in Appendix F (supplementary materials) displays the correlations between total scale and subscale scores on the SOCS-S in this sample.

Internal consistency. Cronbach's alphas ranged from .61 to .90 for total scale and subscale items on the SOCS-O and from .72 to .91 for total scale and subscale items on the SOCS-S (Table 21). These values are acceptable for measures of psychological constructs (Kline, 1999; Terwee et al., 2007).

Floor and ceiling effects. Less than 15% of students received the highest score (100) or lowest score (20) on the SOCS-O and SOCS-S; none received the lowest possible score on the SOCS-O and SOCS-S, and 0% and 0.3% received the highest possible score on the SOCS-O and SOCS-S, respectively. This suggests that both scales capture variability in responses.

Interpretability. Table 22 shows the means and standard deviations of total SOCS-O and SOCS-S scores across subgroups of students. Females scored significantly higher on the SOCS-O compared to males, $t(366) = 2{,}52$, $p = .012$, but there was no significant gender difference in SOCS-S scores, $t(366) = 0.14$, $p = .891$. Length of previous meditation experience had a significant effect on SOCS-S scores only, $F(3) = 4.35$, $p = .005$. Those with 1 to 5 years' meditation experience scored significantly higher on the SOCS-S compared to those without any meditation experience ($p = .019$). Marital status did not significantly affect scores on the SOCS-O ($F(1) = 0.35$, $p = .552$) or SOCS-S ($F(1) = 1.88$, $p = .171$).

Convergent and discriminant validity. Table 23 shows the correlations between total and subscale scores on the SOCS-O and SOCS-S and other constructs. As predicted, the SOCS-O was significantly correlated with the Santa Clara Brief Compassion Scale (SCBCS) and empathic concern (IRI-empathic concern), and the SOCS-S was significantly correlated with the Self-Compassion Scale (SCS-12), at $r \geq .50$. The SOCS-O was also significantly related to the perspective taking (IRI-perspective taking) at $r = .40$ and personal distress (IRI-personal distress) at $r = -.15$. Consistent with expectations, the SOCS-S was significantly correlated in expected

directions with mindfulness (FFMQ-15), wellbeing (SWEMWBS), and stress, anxiety, and depression (DASS), with correlations ranging from moderate-large to large in size. We also found significant small-moderate correlations between the SOCS-O and mindfulness and wellbeing, and a small-moderate, significant, negative correlation between the SOCS-O and depression. Altogether, none of the correlations were $r \geq .80$, at least three quarters of results were consistent with predictions, and at least two correlations were $r \geq .50$, which supports the convergent and discriminant validity of both scales in this student sample.

Relationship between compassion for the self and others. Students scored significantly higher on the SOCS-O compared to the SOCS-S, $t(370) = 19.23$, $p < .001$, $d = 1.15$, 95% CI = 1.01 to 1.29. Table 7 in Appendix F (supplementary materials) presents the correlations between total scale and subscale scores on both measures in the Stage 4 sample. Total SOCS-O and SOCS-S scores were found to significantly correlate at $r = .34$. Many SOCS-O and SOCS-S subscales were also significantly correlated, with coefficients ranging between $r = .03$ and .74. However, the correlation between total scale scores may be artificially inflated given the overlap in universality of suffering items for both scales. We excluded the universality subscale from total SOCS-O and SOCS-S scores and nevertheless found total scores to be significantly correlated at $r = .20$ ($p < .001$).

Discussion

The aim of this programme of research was to develop and evaluate the psychometric properties of two new self-report measures of compassion; the Sussex-Oxford Compassion for Others Scale (SOCS-O) and Sussex-Oxford Compassion for the Self Scale (SOCS-S). These measures were developed and validated across the following four stages: Stage 1 involved consulting experts in contemplative approaches and non-experts to generate and/or review items for each of the theoretically and empirically supported five elements of compassion (Gu et al., 2017; Strauss et al., 2016), Stage 2 applied the a priori five-factor model on data from a large sample of healthcare staff to reduce items for the resulting scales, Stage 3 used CFA in an independent sample of healthcare staff to confirm the factor structures of the SOCS-O

and SOCS-S, and Stage 4 applied CFA on data from University students to cross-validate the factor structures.

Findings from Stages 1 and 2 yielded the 20-item SOCS-O and 20-item SOCS-S and findings from Stages 3 and 4 support the factor structures and demonstrate robust psychometric properties of both scales. For both the SOCS-O and SOCS-S, in both healthcare staff and student samples, a five-factor hierarchical model, with items loading on respective factors from the five-element compassion definition (Strauss et al., 2016) and five factors loading on an overarching compassion factor, was found to fit the data well. All factor intercorrelations and loadings were significant, indicating that the five elements are related and are components of overall compassion for others and compassion for the self constructs. Internal consistency of total SOCS-O and SOCS-S scale and subscale items was adequate in both staff and student samples and the scales showed no indication of floor and ceiling effects. We also facilitated the interpretability of scores on both scales in relation to gender, length of previous meditation experience, and level of education (measured in the healthcare staff sample only). In both healthcare staff and student samples, SOCS-O scores were found to significantly differ between females and males, with females scoring significantly higher compared to males, and in healthcare staff only, SOCS-O scores also significantly differed based on length of previous meditation experience, with those with more previous meditation experience (over 5 years' experience and 1 to 5 years' experience) scoring significantly higher compared to those with little (less than a year's experience) or no meditation experience. In both samples, SOCS-S scores significantly differed based on previous meditation experience, with those with more previous meditation experience scoring significantly higher compared to those with little or no meditation experience, and in healthcare staff, SOCS-S scores also significantly differed based on level of education, with those with higher degrees scoring significantly higher compared to those with GCSEs (UK school qualifications received at age 16) or equivalent qualifications.

The SOCS-O and SOCS-S also showed good evidence of convergent and discriminant validity. Consistent with predictions, in both healthcare staff and student samples, the SOCS-O was significantly correlated with scales measuring compassion for others and empathy and the SOCS-S significantly correlated with an existing self-compassion scale, with correlations large in size, but not so large as to indicate the SOCS-O and SOCS-S are redundant. As hypothesised, the SOCS-S significantly

correlated in expected directions with measures of mindfulness, wellbeing, stress, anxiety, and depression, with correlations ranging from moderate-large to large in size, but not so large as to suggest that they are measuring the same construct. We also found significant small-moderate correlations between the SOCS-O and measures of mindfulness and wellbeing, and significant, small, negative correlations between the SOCS-O and mental health. In the staff sample, both the SOCS-O and SOCS-S were also found to significantly correlate with burnout.

Our findings on the relationship between the SOCS-S and related variables support previous research, but current findings on the SOCS-O contrast with previous research which found no relationship between compassion for others and mindfulness, mental health, wellbeing, and burnout (e.g., Durkin et al., 2016; López et al., 2018; Pommier, 2010). This suggests that the lack of relationship between compassion for others and these constructs may be partly attributable to limitations of existing compassion measures (Strauss et al., 2016). Although we used a cross-sectional design and direction of effects cannot be determined, our findings are consistent with the suggestion that interventions designed to cultivate compassion could improve emotional health outcomes. Future research evaluating the effectiveness of CBIs and other interventions using the SOCS-O and SOCS-S should test this possibility. Similarly, the relationship between compassion and burnout is consistent with the observation of diminishing compassion in cases of work-related burnout in the healthcare sector, although the direction of effect cannot be determined from our findings.

In addition to assessing the relationship between the resulting compassion scales and other measures, we found, for both healthcare staff and students, a significant and small-moderate to moderate correlation between the SOCS-O and SOCS-S. This is at odds with previous research which at best have found small correlations between compassion for the self and others in non-meditator and student samples (López et al., 2018; Neff & Pommier, 2013; Pommier, 2010) and at worst found a small-moderate and negative, but non-significant, correlation between self and other compassion in students (Durkin et al., 2016). Our findings are consistent with the notion that compassion refers to a process that can orient both to the self or others and indicate that self- and other-compassion are overlapping constructs. Previous findings of little or no empirical overlap between the two may be in part due to issues with the measures used in these studies.

Taken together, current findings support the multidimensional conceptualisation of compassion proposed by Strauss et al. (2016) and present the SOCS-O and SOCS-S as new, psychometrically robust self-report measures which overcome limitations of previous compassion scales. We anticipate that these scales will prove valuable in progressing key research areas, including evaluation of the effectiveness of and mechanisms underlying CBIs and further examination of the nature of the relationship between self- and other-compassion.

Limitations and Future Research

The SOCS-O and SOCS-S require further testing; some psychometric properties were not assessed as these were beyond the scope of the current programme of research. These include test-retest reliability, sensitivity to change over the course of a CBI or other interventions which would theoretically cultivate compassion, and further tests of convergent and discriminant validity with additional theoretically related and unrelated constructs.

Although the current programme of research validated the SOCS-O and SOCS-S in healthcare staff and student samples, research in this field has also recruited from other populations (e.g., clinical populations, meditators) and the scales would benefit from cross-validation in such populations to further support their use. Complementing item development through consultation with experts from each of the six continents, future research should also cross-validate the factor structures of the SOCS-O and SOCS-S in samples from other cultures and countries, given that the dominant ethnicity of both healthcare staff and student samples in this study was white and both were UK samples. As part of this line of research, the compassion scales could be translated into different languages which would enable investigation of research questions such as whether there are cross-cultural differences in the strength of the relationship between self- and other-compassion, and compassion and psychological functioning.

Additionally, self-report methods are not without their limitations and are likely to provide only a partial picture of compassion. It would be beneficial for future research to explore whether the SOCS-O and SOCS-S can be triangulated with non-self-report methods of assessing compassion. For example, baseline SOCS-O and SOCS-S scores and/or change in scores over intervention could be correlated with baseline performance and/or change in performance over intervention on behavioural

tasks assessing compassion, such as prosocial games (e.g., the Zurich Prosocial Game; Leiberg, Klimecki, & Singer, 2011) and the recently developed EmpaToM video task, which involves rating compassion levels after watching videos of people recounting emotionally neutral or distressing autobiographical episodes (Kanske, Böckler, Trautwein, & Singer, 2015). Recent findings have found improvements in performance in the Zurich Prosocial Game and EmpaToM task following compassion training (Leiberg et al., 2011; Trautwein, Kanske, Böckler-Raettig, & Singer, 2017) and linked EmpaToM improvements with structural changes in frontoinsular brain regions following compassion training (Valk et al., 2017), another potential marker of increased compassion. However, challenges remain in developing behavioural tasks that can clearly distinguish compassion from distinct but related constructs such as prosocial behaviour, empathy, and altruism. With this in mind, the SOCS-O and SOCS-S have the advantage of accessing the private cognitive and emotional motivations that are part of the compassion construct. They may also be helpful in developing and refining behavioural measures which specifically capture compassion.

Conclusion

Progress in core areas of compassion research requires robust measures that comprehensively capture compassion for others and compassion for the self. The current programme of research developed new theoretically informed and psychometrically robust self-report measures of compassion; the Sussex-Oxford Compassion for Others Scale (SOCS-O) and Sussex-Oxford Compassion for the Self Scale (SOCS-S). Findings support the factor structures of both scales in healthcare staff and student samples. Both the SOCS-O and SOCS-S consist of the following five subscales which can be seen as elements of an overall self- or other-compassion construct: 1) recognising suffering, 2) understanding the universality of suffering in human experience, 3) feeling for the person suffering and emotionally connecting with their distress, 4) tolerating uncomfortable feelings aroused so that we remain open to and accepting of them in their suffering, and 5) acting or being motivated to act to alleviate suffering. Findings also support the psychometric properties of both scales in terms of their internal consistency, interpretability, floor and ceiling effects, and convergent and discriminant validity. Taken together, the rigorous development process employed in the current research programme and emergent psychometric

properties of the SOCS-O and SOCS-S support their use in compassion research and practice.

Chapter 8:
General Discussion

Summary of Findings

The aim of this thesis was to address several important omissions in the mindfulness and compassion research literature, concerning the mechanisms of mindfulness-based interventions (MBIs) and measurement of mindfulness, and the lack of definitional and measurement clarity of compassion. Addressing these gaps not only furthers research and understanding of these contemplative constructs but is crucial to building a robust evidence base in these fields and increasing the public health impact of MBIs and compassion-based interventions (CBIs).

Chapter 2 consisted of a systematic review and meta-analysis of studies which formally tested the mechanisms of mindfulness-based stress reduction (MBSR) and mindfulness-based cognitive therapy (MBCT) using mediation analysis. The narrative synthesis of findings across the 20 included studies identified moderate and consistent evidence for mindfulness, rumination, and worry as mechanisms of MBSR/MBCT, and preliminary but insufficient evidence for self-compassion, cognitive and emotional reactivity, and psychological flexibility as mechanisms of action. Two-stage structural equation modeling (TSSEM) analyses additionally found both mindfulness and repetitive negative thinking (worry and rumination) to be significant mediators of the effect of MBIs on mental health outcomes (anxiety, depression, global psychopathological symptoms, stress, negative affect). This review also assessed the quality of included studies and noted that because most studies have key methodological limitations (e.g., lack of matched control conditions and a general failure to measure mediators prior to outcomes), this prevents robust conclusions regarding mediation.

Chapter 3 examined the specific effects and mechanisms of learning mindfulness on stress in a sample of University students and staff, by comparing a two-week online mindfulness-based self-help (MBSH) intervention with a well-matched control condition and waitlist control condition. This study found that compared to both control conditions, MBSH significantly reduced stress at post-intervention. The matched control condition, listening to classical music, was rated by participants as being as equally plausible and engaging as MBSH, which suggests that MBSH is an effective way of reducing stress in the student and staff population that is not simply attributable to a placebo effect. Bootstrapping-based mediation analyses showed that changes in mindfulness, self-compassion, and worry significantly mediated the effects

of MBSH compared to both control conditions on changes in stress. This supports the suggestion that mindfulness, self-compassion, and worry are mechanisms of action which are specific to self-help MBIs, rather than general processes which underlie change in any plausible self-help intervention.

Chapter 4 used confirmatory factor analysis (CFA) to assess the stability of the factor structure of the commonly used Five-Facet Mindfulness Questionnaire (FFMQ) before and after MBCT in a sample of adults with recurrent depression in remission. Before MBCT, findings showed that a four-factor hierarchical model (excluding the 'observing' facet) best fit the FFMQ data, whereas after MBCT, a five-factor hierarchical model best represented the data. These findings support the notion that meditation experience changes people's qualities of noticing and suggest that the observing subscale should be excluded from comparisons of total scale and subscale scores before and after mindfulness interventions. This chapter also assessed the factor structure and psychometric properties of a shortened 15-item version of the FFMQ and findings support the use of this version.

Chapter 5 presented a review of definitions of compassion and a systematic review and evaluation of compassion questionnaire measures. Existing conceptualisations of compassion were consolidated into the following five-element definition, which can apply to both compassion for the self and for others; compassion consists of recognising suffering, understanding the universality of suffering in human experience, feeling for the person suffering, tolerating uncomfortable feelings, and motivation to act or acting to alleviate suffering. The systematic review identified nine existing compassion measures, none of which were deemed to comprehensively capture compassion or have strong psychometric properties. Findings called for empirical testing of the five-element definition, and if this is supported, the development of self- and other-compassion measures which are based on this operational definition and demonstrates adequate psychometric properties.

Chapter 6 empirically investigated the conceptual structure of the five-element definition of compassion by conducting exploratory factor analysis (EFA) and CFA of self-report items. Findings from the EFA supported a five-factor structure of compassion consistent with the five-element definition and findings from the CFA also showed promising support for this five-factor model. A number of item-level limitations indicate that items from included measures should not be used to

collectively assess compassion. Instead, the five-element definition should be used to guide item generation for the development of a new self-report measure of compassion.

Chapter 7 used the five-element definition as a basis for the development and validation of new self-report measures of compassion for the self and for others. Findings from the item generation, review, and reduction stages yielded the 20-item Sussex-Oxford Compassion for Others Scale (SOCS-O) and 20-item Sussex-Oxford Compassion for the Self Scale (SOCS-S). CFA findings support the five-factor structure from the five-element definition for both scales. Both measures were also found to demonstrate robust psychometric properties in terms of internal consistency, interpretability, floor and ceiling effects, and convergent and discriminant validity.

Strengths of the Research Programme

The strengths specific to each study have been detailed in individual chapters (Chapters 2 to 7). This section instead takes a broader perspective and discusses the general strengths of the programme of research undertaken in this thesis.

A key strength of the current research programme is that it yields novel contributions to the literature on mindfulness and compassion, and such contributions are directed towards addressing key and timely issues necessary for the continued development of both fields. In their comprehensive review mapping the evidence base for MBIs, Dimidjian and Segal (2015) highlighted areas of saturation, in terms of the proportional amount of published research on MBIs, and important gaps that warrant greater attention. They found that the greatest focus of activity in the field has been focused on generating and refining MBIs and evaluating their effectiveness in research settings. The authors note that if clinical and research attention remain dedicated to these areas, this is likely to limit the public health impact of MBIs. By contrast, Dimidjian and Segal highlight that fewer studies have attended to the 'basics' in terms of clarifying mediating processes of change (i.e., *how* do MBIs work?) and intervention targets (i.e., *what* outcomes do MBIs affect?). This work is needed to provide a scientific foundation for how and why an intervention may be helpful for a particular population and problem. A focus of the current thesis has been on clarifying the mechanisms of MBIs, therefore addressing one of the key gaps in the field.

Attending to the quality of the tools with which we measure key constructs is also critical to scientific growth. This basic area of research can be overlooked in the enthusiasm to develop and evaluate novel interventions which follows early research indicating benefits of enhancing mindfulness and compassion. However, continued use of measures which lack validity and reliability can undermine research efforts in these areas, as valid and reliable measures of mindfulness and compassion are needed before we can fully examine effectiveness and mechanisms of change. Another focus of the current thesis has been on improving and clarifying the measurement of mindfulness and compassion, with the intention of strengthening the foundation of both fields.

The studies which comprise this thesis addressed the aforementioned research gaps using a range of samples (e.g., NHS staff, University students and staff, people with recurrent major depressive disorder) and multiple methods (e.g., systematic reviews, meta-analysis, mediation analysis, randomised controlled trial, measurement development and psychometric testing), with individual studies using samples and methods appropriate to their aim and design. The deliberate inclusion of different samples and/or methods to address each key omission around measurement and mechanism broadens the scope of the current findings.

A further strength is the attention and emphasis given throughout the current programme of research on methodological rigor. For example, Chapter 2 evaluated the quality of mediation studies included in the systematic review and concluded that methodological limitations of included studies prevent robust conclusions regarding the mechanisms of MBIs. Chapter 3 improved on methodological limitations of many MBSH studies, by including a well-matched control condition, investigating theoretically and empirically supported mechanisms, and using recommended approaches to mediation analysis (bootstrapping, multicategorical IV mediation models). Chapters 4 to 7 were built on the need for psychometrically robust measures of mindfulness and compassion. The new SOCS-O and SOCS-S were developed and evaluated according to good practice guidelines and with the limitations of previous compassion measures in mind.

Limitations and Future Directions

The specific limitations of each study have been detailed in individual chapters (Chapters 2 to 7). This section reflects on points for consideration arising from the programme of research as a whole. General directions for future research are discussed in light of these considerations.

Although the current programme of research used multiple methods to address omissions in measurement and mechanism research, mindfulness and compassion outcomes were assessed using self-report questionnaire measures only and all methods employed were quantitative in approach. Chapter 1 included justifications for the use of self-report measures of mindfulness and compassion and appropriate quantitative analyses were selected to address particular research questions. Nevertheless, it is important to explore the development and validation of behavioural and objective methods of measuring mindfulness and compassion, and triangulate multiple methods of assessment. Exploring non-self-report measures has the potential to facilitate further understanding of these contemplative constructs and improve how they are measured. Research on the mechanisms of MBIs would also benefit from a mixed-methods approach, where mediation analyses on quantitative data are complemented by qualitative data from participants, detailing their experiences of change following interventions. This has the potential to yield rich data regarding change processes and how these may develop and interact over the course of an MBI. It would be particularly interesting to collect both quantitative and qualitative data from a single sample and compare the mechanisms arising from each set of findings. Such qualitative data could also be used to develop new theoretical models of the mechanisms of MBIs or to refine existing models.

All of the studies in this thesis used measures of mindfulness and/or compassion which conceptualise these constructs as dispositional traits, consistent across time and contexts, rather than as states, with capacities fluctuating across time and situations. In terms of developing and refining MBIs and CBIs, and examining mechanisms and trajectories of change, it would also be beneficial to use measures of mindfulness and compassion which assess these as state-like constructs. This would be particularly valuable given the assumption of MBIs and CBIs that a dispositional tendency to be mindful and compassionate is developed through repeated practice of skills that cultivate mindful and compassionate states. Research using state measures may also shed light on the facilitators and barriers to cultivating mindfulness and compassion in particular contexts. Currently, self-report measures of state mindfulness exist (e.g.,

State Mindfulness Scale; Tanay & Bernstein, 2013, and Toronto Mindfulness Scale; Lau et al., 2006) but there are no measures of state compassion. Strauss et al. (2016) stated that their five-element definition can be applied to both state and trait compassion and future research could use this definition to develop a measure of state compassion.

Moreover, Chapters 5 to 7 focused on examining and measuring compassion at the level of the individual. This may present an incomplete picture of compassion, implying that it is solely the responsibility of the individual and disregarding the role of external circumstances. This is particularly important to consider in relation to the healthcare sector, where care and compassion are viewed as essential qualities of the workforce, and blame may fall solely on individual healthcare professionals for failings in compassionate care that could be attributed to external, organisation-level factors, such as resourcing constraints or restructuring. Future research should examine how individual levels of compassion interact with external factors and the conditions under which an individual's experience of compassion may be facilitated or hindered. Such research could employ a mixed-methods approach, using a measure of state compassion and examining the emergence of compassion under various experimental conditions, and exploring experiences of compassion in different contexts using qualitative methods. Potential advantages of such research include informing the development of CBIs and organisation-level changes to create work environments which best support the emergence of compassion.

Priorities for Future Research

In addition to the general directions for future research summarised in the previous section, this section outlines key areas future research should prioritise, based on the findings from individual studies in this thesis.

An unexpected but key finding from the systematic review of studies which examined the mechanisms of MBIs (Chapter 2) is that most reviewed studies have a moderate risk of bias and at least one important methodological shortcoming. For example, many studies lacked active control conditions, did not establish whether change in mediators occurred prior to change in outcomes, and did not use recommended methods of mediation analysis. The priority for future research

examining the mechanisms of MBIs should be to improve upon these limitations, so that we can make stronger conclusions regarding how MBIs work. Once we are confident in the evidence base, future research can begin to broaden the study of mechanisms of MBIs, exploring for example, whether they differ for particular populations and problems and how multiple mechanisms interact to produce change.

Findings from Chapter 4 showed that the factor structure of the FFMQ is not invariant before and after MBCT, which informs the use of this popular measure. More research is required to determine whether these findings are replicated in other clinical and non-clinical samples and using other MBIs (e.g., MBSR) and interventions with less meditation practice (e.g., MBSH interventions). Whether or not these findings are replicated in future work would contribute to the continued evaluation and use of the FFMQ and existing theory that meditation experience alters the way in which we observe phenomena. Additional psychometric testing of the FFMQ-15 is also required, including tests of discriminant validity and test-retest reliability.

The SOCS-O and SOCS-S developed in Chapter 7 were found to demonstrate robust psychometric properties but require further evaluation. Further psychometric testing needed include sensitivity to change over the course of a CBI, MBI, or another type of intervention which would theoretically increase compassion, test-retest reliability, cross-validation of the factor structure in other populations (e.g., meditators, clinical populations), and further tests of convergent and discriminant validity with additional constructs. Continued assessment of these scales would be valuable in terms of further supporting their use, highlighting any areas for refinement to improve their use, and increasing understanding of the core features of compassion and how they interact.

Research Implications

Current findings contribute to our understanding of how mindfulness interventions (MBIs and MBSH interventions) may improve psychological functioning. They indicate support for increased mindfulness and self-compassion and decreased repetitive negative thinking (worry and rumination) as mechanisms of action. Findings also have implications for future mechanism research. They indicate that based on the quality of existing evidence, at present, conclusions regarding these constructs as

mechanisms should be tentative. Future work should therefore aim to improve on the methodological limitations of the current body of evidence. If support for mindfulness, repetitive negative thinking, and self-compassion is replicated in future mediation studies, this would strengthen their position as mechanisms of MBIs. This would also have clinical implications and suggest that MBIs may be optimised by emphasising teaching and practice related to mindfulness, rumination, worry, and self-compassion. In addition, if future work continues to show an overlap between the mechanisms underlying MBIs and MBSH interventions, this would strengthen the suggestion that the latter may be promising alternatives to MBIs in particular situations.

Findings from the current research programme also have theoretical and practical implications for our understanding of the construct of mindfulness and use of the FFMQ, respectively. The key finding from Chapter 4 was that before MBCT, a four-factor structure of the FFMQ, which excludes the observing subscale, provided the best fit, but after MBCT, a five-factor structure with all five facets best fit the data. This, considered alongside studies demonstrating different factor structures of the FFMQ (with and without observing) in meditator and non-meditator samples (e.g., Baer et al., 2008, M. J. Williams et al., 2014), supports the perspective that meditation experience changes the way in which people notice phenomena. For those who have not undertaken an MBI and have little or no meditation experience, the way in which they observe internal and external events may not be consistent with mindful qualities (i.e., the *how* aspect of mindfulness). Following an MBI and experience of meditation, the way in which these same people notice may be more consistent with mindful awareness. In addition to informing our understanding of how MBIs may alter mindful awareness, current findings showing support for hierarchical models of the FFMQ strengthen the theoretical conceptualisation of mindfulness as being a multifaceted yet coherent construct.

In terms of implications for our use of the FFMQ in research and clinical practice, findings that meditation experience alters people's qualities of observing indicate that pre- and post-MBI comparisons of total FFMQ scores and scores on the observing subscale may not be valid. These findings suggest that the observing facet should be omitted when comparing total scale and subscale scores before and after MBIs. Findings from Chapter 4 also validated a short, 15-item version of the FFMQ, and support its use as an alternative measure to the original FFMQ. This version is anticipated to be helpful in research where briefer forms are needed.

Finally, current findings have theoretical implications for our understanding of compassion constructs, and how they relate to each other and to psychological outcomes, and practical implications for how we measure compassion. The review of compassion conceptualisations in Chapter 5 culminated in the proposal of a comprehensive five-element definition of compassion for the self and compassion for others. This review contributed to greater clarity on the key features which comprise compassion and the definition provided a necessary foundation for first evaluating the content validity of existing compassion measures and later developing new measures of compassion when existing ones were found to be lacking.

The SOCS-O and SOCS-S, based on the five-element definition of compassion, were found in the current programme of research to be psychometrically robust measures of compassion and overcome many of the limitations of previous scales. Being able to measure compassion using valid and reliable tools is necessary for the growth of this field. In addition to this contribution, psychometric evaluation of both scales informs our understanding of how compassion relates to other constructs and the nature of the relationship between self- and other-compassion. Chapter 7 found that greater compassion for the self and for others were related to increased mindfulness and wellbeing, and decreased burnout, stress, depression, and anxiety. This supports the cultivation of compassion to improve psychological functioning. These findings are particularly valuable given that previous studies have failed to find links between compassion for others and mindfulness, wellbeing, and mental health outcomes (e.g., Durkin et al., 2016; López et al., 2018; Pommier, 2010). This suggests that previous findings may be in part attributable to limitations of existing compassion measures.

Chapter 7 also showed that the same five-factor model fit both self-compassion data and other-compassion data well, and found small-moderate to moderate, and significant, correlations between the SOCS-O and SOCS-S. This contributes to the discourse on the relationship between self- and other-compassion and has implications for future research examining this association. Current findings suggest that the components that make up the overall experience of compassion may be the same irrespective of target, and that there is an overlap in the degree to which we experience elements of compassion towards ourselves and other people. This contrasts to previous research which found little or no empirical overlap between compassion for the self and others (Durkin et al., 2016, López et al., 2018; Neff & Pommier, 2013; Pommier, 2010). Previous findings may be partly due to issues with the compassion measures

used in these studies. Future research should therefore not be deterred by initial findings indicating no relationship between self- and other-compassion. Current findings open up possibilities for further research, including whether there are differences cross-culturally, and between meditators and non-meditators, in the strength of the relationship between self- and other-compassion, and whether compassion for the self is supportive of compassion for others.

Clinical Implications

The current body of work also has broader clinical implications. Given the prevalence and negative personal and economic consequences of poor mental health and wellbeing (e.g., Chartered Institute of Personnel and Development, 2016; Hassard et al., 2018; Health and Safety Executive, 2017; McCrone et al., 2008; World Health Organisation, 2017), identifying effective treatments is a priority. Evidence shows that MBIs are effective in treating depression (both in reducing relative risk of relapse and in treating symptoms of a current episode) and for improving stress and wellbeing (Khoury et al., 2015; Kuyken et al., 2016; Strauss et al., 2014), indicating that they can be part of the solution to poor mental health and wellbeing. CBIs also show promise as treatments for depression, anxiety, wellbeing, and psychological distress (Kirby et al., 2017).

Findings from the current research programme address significant omissions in mindfulness and compassion research and make important contributions to the following key areas: 1) improving measurement and understanding of the construct of mindfulness, 2) clarifying the definition of compassion, 3) providing valid and reliable measures of compassion, and 4) clarifying and empirically testing the mechanisms of MBIs, in order to determine whether interventions work through proposed mechanisms and that findings are not simply a placebo effect. Demonstrating that improvements following MBI are not a placebo effect would help inform the use of limited public health money and support the implementation of MBIs. Moreover, by attending to these crucial yet underdeveloped areas of mindfulness and compassion research, this thesis contributes to building a robust evidence base for MBIs and CBIs and as a consequence, increasing their potential for public health impact.

Conclusions

Despite the growing research interest in mindfulness and compassion, and their potential to improve mental health and wellbeing, there remain significant research gaps that limit progress in these fields. Important gaps include clarifying the definition of compassion, measurement of mindfulness and compassion, and mechanisms of MBIs. The aim of this thesis was to address these omissions.

Findings from Chapters 2 and 3 inform our understanding of how mindfulness interventions may improve psychological functioning; they support mindfulness, worry, rumination, and self-compassion as mechanisms of change. They also highlight priorities for future research: to improve upon the key methodological limitations of many studies which investigated mechanisms, so that stronger conclusions can be made regarding how mindfulness interventions work. Findings from Chapter 4 developed our understanding of how the experience of mindfulness may change over the course of an MBI. They also have practical implications for the measurement of mindfulness in effectiveness and mechanism research; they suggest that the observing subscale of the widely-used FFMQ should be excluded from comparisons of total scale and subscale scores before and after mindfulness interventions.

In an attempt to clarify the definition of compassion and identify key elements which comprise compassion, Chapter 5 reviewed and consolidated existing theoretical conceptualisations into a five-element definition. Another key contribution from this chapter was the finding that no existing compassion measures comprehensively capture compassion or have strong psychometric properties, which necessitates the development of new compassion measures. Findings from Chapter 6 empirically supported the conceptual structure of the five-element definition and provided additional support for the use of this definition as the basis for the development of the SOCS-O and SOCS-S in Chapter 7.

Taken together, findings from the current programme of research addressed significant gaps and provide novel contributions to the literature on mindfulness and compassion. They yield insights into the nature of mindfulness and compassion, construct necessary foundations for further work, and highlight important future directions.

References

References preceded by one asterisk (*) indicate studies included in the systematic review in Chapter 2. References preceded by two asterisks (**) indicate measures included in the systematic review in Chapter 5.

Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

American Medical Association. (2001). Principles of medical ethics. Retrieved from www.ama-assn.org/sites/default/files/media-browser/principles-of-medical-ethics.pdf

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th Ed.). Washington, DC: Author.

Ardelt, M. (2003). Empirical Assessment of a Three-Dimensional Wisdom Scale. *Research on Aging, 25*, 275–324.

Baer, R. A. (2003). Mindfulness training as a clinical intervention: A conceptual and empirical review. *Clinical Psychology: Science and Practice, 10*(2), 125-143.

Baer, R. A. (2011). Measuring mindfulness. *Contemporary Buddhism, 12*(01), 241-261.

Baer, R. (2015). Ethics, values, virtues, and character strengths in mindfulness-based interventions: A psychological science perspective. *Mindfulness, 6*(4), 956-969.

Baer, R. A., Carmody, J., & Hunsinger, M. (2012). Weekly change in mindfulness and perceived stress in a mindfulness-based stress reduction program. *Journal of Clinical Psychology, 68*(7), 755-765.

Baer, R. A., Gu, J., Cavanagh, K., & Strauss, C. Discriminant validity of mindfulness questionnaires: A meta-analysis of randomised controlled trials comparing mindfulness-based interventions to active control conditions on self-report measures of mindfulness. *Manuscript under review.*

Baer, R. A., Lykins, E. L. B., & Peters, J. R. (2012). Mindfulness and self-compassion as predictors of psychological wellbeing in long-term meditators and matched nonmeditators. *The Journal of Positive Psychology, 7*(3), 230-238.

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-

report: The Kentucky Inventory of Mindfulness Skills. *Assessment, 11*, 191–206.

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment, 13*(1), 27-45.

Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., . . . Williams, J. M. G. (2008). Construct validity of the Five Facet Mindfulness Questionnaire in meditating and nonmeditating samples. *Assessment, 15*, 329–342.

Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In: G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 269–296). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Barker, C., Pistrang, N., & Elliott, R. (2002). *Research methods in clinical psychology*. London: Wiley.

Barnard, P. J., Watkins, E. R., Mackintosh, B., & Nimmo–Smith, I. (2007). *Getting Stuck in a Mental Rut: some process and experiential attributes*. Paper presented at the BABCP Annual Conference, Brighton, 13–15 September.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Barrett, P. (2005). Shortform 1.1 [Software]. Available from www.pbarrett.net/#software.

*Batink, T., Peeters, F., Geschwind, N., van Os, J., & Wichers, M. (2013). How does MBCT for depression work? Studying cognitive and affective mediation pathways. *PloS one, 8*(8), e72778.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588-606.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling.

Sociological Methods & Research, 16(1), 78-117.

*Bergen-Cico, D., & Cheon, S. (2013). The mediating effects of mindfulness and self-compassion on trait anxiety. *Mindfulness, 5*(5), 505-519.

Bernard, P., Cullen, C., & Kuyken, W. (2017). *MBCT for life teacher manual.* Oxford: Oxford Mindfulness Centre.

Bhikkhu, T. (2007). *Mindfulness defined*. Retrieved from www.dhammatalks.org/Archive/Writings/CrossIndexed/Uncollected/MiscEssays/Mindfulness%20Defined.pdf

*Bieling, P. J., Hawley, L. L., Bloch, R. T., Corcoran, K. M., Levitan, R. D., Young, L. T., . . . Segal, Z. V. (2012). Treatment specific changes in decentering following mindfulness-based cognitive therapy versus antidepressant medication or placebo for prevention of depressive relapse. *Journal of Consulting and Clinical Psychology, 80*(3), 365-372.

Birnie, K., Speca, M., & Carlson, L. E. (2010). Exploring self-compassion and empathy in the context of mindfulness-based stress reduction (MBSR). *Stress and Health, 26*(5), 359-371.

Bishop, S. R. (2002). What do we really know about mindfulness-based stress reduction? *Psychosomatic medicine, 64*(1), 71-83.

Bohlmeijer, E., Prenger, R., Taal, E., & Cuijpers, P. (2010). The effects of mindfulness-based stress reduction therapy on mental health of adults with a chronic medical disease: a meta-analysis. *Journal of psychosomatic research, 68*(6), 539-544.

Bohlmeijer, E., ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment, 18*(3), 308-320.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research, 21*, 205–229.

Bowen, S., Chawla, N., & Marlatt, G. A. (2010). *Mindfulness-based relapse prevention for addictive behaviors: A clinician's guide*. London, UK: Guilford Press.

*Bränström, R., Kvillemo, P., Brandberg, Y., & Moskowitz, J. T. (2010). Self-report mindfulness as a mediator of psychological well-being in a stress reduction intervention for cancer patients—A randomized study. *Annals of Behavioral Medicine, 39*(2), 151-161.

Brittlebank, A. D., Scott, J., Williams, J. M. G., & Ferrier, I. N. (1993).

Autobiographical memory in depression: State or trait marker. *British Journal of Psychiatry, 162*, 118 –121.

*Britton, W. B., Shahar, B., Szepsenwol, O., & Jacobs, W. J. (2012). Mindfulness-based cognitive therapy improves emotional reactivity to social stress: Results from a randomized controlled trial. *Behavior Therapy, 43*(2), 365-380.

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*, 822–848.

Brown, K. W., Ryan, R. M., & Creswell, J. D. (2007). Mindfulness: Theoretical foundations and evidence for its salutary effects. *Psychological Inquiry, 18*(4), 211-237.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In: K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Beverly Hills, CA: Sage.

**Burnell, L., & Agan, D. L. (2013). Compassionate care: Can it be defined and measured? The development of the Compassionate Care Assessment Tool. *Int J Soc Res Methodol, 6*(2), 180-7.

Burns, J. L., Labbé, E., Arke, B., Capeless, K., Cooksey, B., Steadman, A., & Gonzales, C. (2002). The effects of different types of music on perceived and physiological measures of stress. *Journal of Music Therapy, 39*(2), 101-116.

Byrne, B. M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of personality assessment, 85*(1), 17-32.

Cameron, R. A., Mazer, B. L., DeLuca, J. M., Mohile, S. G., & Epstein, R. M. (2015). In search of compassion: A new taxonomy of compassionate physician behaviours. *Health Expectations: An International Journal of Public Participation in Health Care and Health Policy, 18*(5), 1672-1685.

Carissoli, C., Villani, D., & Riva, G. (2015). Does a meditation protocol supported by a mobile application help people reduce stress? Suggestions from a controlled pragmatic trial. *Cyberpsychology, Behavior, and Social Networking, 18*(1), 46-53.

Carmody, J., & Baer, R. A. (2008). Relationships between mindfulness practice and

levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. *Journal of Behavioral Medicine, 31*(1), 23-33.

Cassell, E. J. (2002). Compassion. In C. R. Synder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 434-445). Oxford: Oxford University Press.

Cavanagh, K., Strauss, C., Cicconi, F., Griffiths, N., Wyper, A., & Jones, F. (2013). A randomised controlled trial of a brief online mindfulness-based intervention. *Behaviour Research and Therapy, 51*(9), 573-578.

Cavanagh, K., Strauss, C., Forder, L., & Jones, F. (2014). Can mindfulness and acceptance be learnt by self-help? A systematic review and meta-analysis of mindfulness and acceptance-based self-help interventions. *Clinical Psychology Review, 34*(2), 118-129.

Chadwick, P., Hember, M., Symes, J., Peters, E., Kuipers, E., & Dagnan, D. (2008). Responding mindfully to unpleasant thoughts and images: Reliability and validity of the southampton mindfulness questionnaire (SMQ). *British Journal of Clinical Psychology, 47*(4), 451-455.

Chartered Institute of Personnel and Development. (2016). *Growing the health and well-being agenda: From first steps to full potential*. Retrieved from www.cipd.co.uk/Images/health-well-being-agenda_2016-first-steps-full-potential_tcm18-10453.pdf

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological methods & research, 36*(4), 462-494.

Cheung, M. W. L. (2008). A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. *Psychological Methods, 13*(3), 182-202.

Cheung, M. W. L. (2013). metaSEM: Meta-analysis using structural equation modeling. R package version 0.8-4. Retrieved from courses.nus.edu.sg/course/psycwlm/Internet/metaSEM/

Cheung, M. W. L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods, 10*(1), 40-64.

Cheung, M. W. L., & Cheung, S. F. (2016). Random-effects models for meta-analytic structural equation modeling: review, issues, and illustrations. *Research Synthesis Methods, 7*(2), 140-155.

Chiesa, A., Anselmi, R., & Serretti, A. (2014). Psychological mechanisms of mindfulness-based interventions: What do we know? *Holistic Nursing Practice, 28*(2), 124-148.

Chiesa, A., & Malinowski, P. (2011). Mindfulness-based approaches: Are they all the same? *Journal of Clinical Psychology, 67*(4), 404-424.

Chiesa, A., & Serretti, A. (2009). Mindfulness-based stress reduction for stress management in healthy people: A review and meta-analysis. *The Journal of Alternative and Complementary Medicine, 15*(5), 593-600.

Chou, C. P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In: R. H. Hoyle (Eds.), *Structural equation modeling: Issues and applications* (pp. 37–55). Newbury Park, CA: Sage.

Christopher, M. S., Neuser, N. J., Michael, P. G., & Baitmangalkar, A. (2012). Exploring the psychometric properties of the five facet mindfulness questionnaire. *Mindfulness, 3*(2), 124-131.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*(4), 385-396.

Compassion in Education Foundation. (2014). Bringing compassion into education and learning: The Second Strategic Plan & Progress Report. Retrieved from www.coedfoundation.org.uk/pdfs/CoEd_ Foundation_strategic_plan.pdf

Compassion in Education Foundation. (2016). Bringing compassion into education and learning: The Third Strategic Plan & Progress Report. Retrieved from www.coedfoundation.org.uk/pdfs/CoEd_ Foundation_strategic_plan.pdf

Cooper, H., & Patall, E. A. (2009). The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological Methods, 14*(2), 165-176.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *The Journal of Applied Psychology, 78*(1), 98-104.

Cosley, B. J., McCoy, S. K., Saslow, L. R., & Epel, E. S. (2010). Is compassion for others stress buffering? Consequences of compassion and social support for physiological reactivity to stress. *Journal of Experimental Social Psychology, 46*, 816–823.

Costa, J., Marôco, J., Pinto-Gouveia, J., Ferreira, C., & Castilho, P. (2015). Validation of the psychometric properties of the self-compassion scale. testing the factorial validity and factorial invariance of the measure among borderline personality

disorder, anxiety disorder, eating disorder and general populations. *Clinical Psychology & Psychotherapy, 23*(5), 460-468.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research & evaluation, 10*(7), 1-9.

Crane, C., Crane, R. S., Eames, C., Fennell, M. J., Silverton, S., Williams, J. M. G., & Barnhofer, T. (2014). The effects of amount of home meditation practice in Mindfulness Based Cognitive Therapy on hazard of relapse to depression in the Staying Well after Depression Trial. *Behaviour research and therapy, 63,* 17-24.

Crane, R. S., Eames, C., Kuyken, W., Hastings, R. P., Williams, J. M., Bartley, T., . . . Surawy, C. (2013). Development and validation of the mindfulness-based interventions—Teaching assessment criteria (MBI:TAC). *Assessment, 20*(6), 681-688.

*Crane, C., Winder, R., Hargus, E., Amarasinghe, M., & Barnhofer, T. (2012). Effects of mindfulness-based cognitive therapy on specificity of life goals. *Cognitive Therapy and Research, 36*(3), 182-189.

Crawford, P., Brown, B., Kvangarsnes, M., & Gilbert, P. (2014). The design of compassionate care. *Journal of Clinical Nursing, 23*(23-24), 3589-3599.

Crocker, J., & Canevello, A. (2008). Creating and undermining social support in communal relationships: the role of compassionate and self-image goals. *Journal of personality and social psychology, 95*(3), 555.

Cuijpers, P., Donker, T., van Straten, A., Li, J., & Andersson, G. (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine, 40*(12), 1943-1957.

Curtiss, J., & Klemanski, D. H. (2014a). Factor analysis of the five facet mindfulness questionnaire in a heterogeneous clinical sample. *Journal of Psychopathology and Behavioral Assessment, 36*(4), 683-694.

Curtiss, J., & Klemanski, D. H. (2014b). Teasing apart low mindfulness: Differentiating deficits in mindfulness and in psychological flexibility in predicting symptoms of generalized anxiety disorder and depression. *Journal of Affective Disorders, 166*, 41-47.

Dahlsgaard, K., Peterson, C., & Seligman, M. E. (2005). Shared virtue: The

convergence of valued human strengths across culture and history. *Review of general psychology, 9*, 203.

Dalai Lama. (1995). *The power of compassion*. New Delhi: Harper Collins.

Dalai Lama. (2000). Training the mind: Verse 7. In: *Eight verses for training the mind (Verse 7)*. Retrieved from www.dalailama.com/teachings/training-the-mind/training-the-mind-verse-7

Dalai Lama. (2002a). Dialogues, part II: Pragmatic extensions and applications. In: R. Davidson & A. Harrington (Eds.), *Visions of compassion: Western scientists and Tibetan Buddhists examine human nature* (pp. 213-246). Oxford: Oxford University Press.

Dalai Lama. (2002b). *How to practice: The way to a meaningful life.* New York: Pocket Books.

Dalai Lama. (2005). *Essence of the Heart Sutra: The Dalai Lama's Heart of Wisdom Teachings.* USA: Wisdom Publications.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.

Davidson, R. J. & Schuyler, B. S. (2015). Neuroscience of happiness. In: J. F. Helliwell, R. Layard, & J. Sachs (Eds.), *World Happiness Report 2015* (pp. 88-106). New York: Sustainable Development Solutions Network.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.

de Waal, F. (2009). *The age of empathy*. New York: Harmony.

Department of Health. (2013). The handbook to the NHS constitution. Retrieved from www.gov.uk/government/publications/the-nhs-constitution-for-england

Desbordes, G., Negi, L. T., Pace, T. W., Wallace, B. A., Raison, C. L., & Schwartz, E. L. (2012). Effects of mindful-attention and compassion meditation training on amygdala response to emotional stimuli in an ordinary, non-meditative state. *Frontiers in human neuroscience, 6*, 292.

Desrosiers, A., Vine, V., Curtiss, J., & Klemanski, D. H. (2014). Observing

nonreactively: A conditional process model linking mindfulness facets, cognitive emotion regulation strategies, and depression and anxiety symptoms. *Journal of Affective Disorders, 165*, 31-37.

DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26). Sage publications.

Dewar, B., Pullin, S., & Tocheris, R. (2011). Valuing compassion through definition and measurement. *Nursing Management, 17*(9), 32-7.

Dimidjian, S., Beck, A., Felder, J. N., Boggs, J. M., Gallop, R., & Segal, Z. V. (2014). Web-based mindfulness-based cognitive therapy for reducing residual depressive symptoms: an open trial and quasi-experimental comparison to propensity score matched controls. *Behaviour research and therapy*, *63*, 83-89.

Dimidjian, S., & Segal, Z. V. (2015). Prospects for a clinical science of mindfulness-based intervention. *American Psychologist, 70*(7), 593.

Dobkin, P. L., Irving, J. A., & Amar, S. (2012). For whom may participation in a mindfulness-based stress reduction program be contraindicated? *Mindfulness*, *3*(1), 44-50.

Dowd, H., Hogan, M. J., McGuire, B. E., Davis, M. C., Sarma, K. M., Fish, R. A., & Zautra, A. J. (2015). Comparison of an online mindfulness-based cognitive therapy intervention with online pain management psychoeducation: a randomized controlled study. *The Clinical journal of pain, 31*(6), 517-527.

Dreher, D. E., Holloway, K. A., & Schoenfelder, E. (2007). The vocation identity questionnaire: Measuring the sense of calling. *Research in the Social Scientific Study of Religion, 18*, 99.

Durkin, M., Beaumont, E., Martin, C. J. H., & Carson, J. (2016). A pilot study exploring the relationship between self-compassion, self-judgement, self-kindness, compassion, professional quality of life and wellbeing among UK community nurses. *Nurse education today, 46*, 109-114.

Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy, 1*(3), 192-205.

Ehring, T., Zetsche, U., Weidacker, K., Wahl, K., Schonfeld, S., & Ehlers, A. (2011). The perseverative thinking questionnaire (PTQ): Validation of a content-independent measure of repetitive negative thinking. *Journal of Behavior Therapy and Experimental Psychiatry, 42*(2), 225-232.

Ekman, P. (2016). What scientists who study emotion agree about. *Perspectives on*

Psychological Science, 11(1), 31-34.

Eisenberg, N., Guthrie, I. K., Cumberland, A., Murphy, B. C., Shepard, S. A., Zhou, Q., & Carlo, G. (2002). Prosocial development in early adulthood: A longitudinal study. *Journal of Personality and Social Psychology, 82*(6), 993-1006.

Eisenberg, N., VanSchyndel, S. K., & Hofer, C. (2015). The association of maternal socialization in childhood and adolescence with adult offsprings' sympathy/caring. *Developmental Psychology, 51*(1), 7-16.

Epstein, R. M., Franks, P., Shields, C. G., Meldrum, S. C., Miller, K. N., Campbell, T. L., & Fiscella, K. (2005). Patient-centered communication and diagnostic testing. *The Annals of Family Medicine, 3*(5), 415–421.

Feldman, G., Hayes, A., Kumar, S., Greeson, J., & Laurenceau, J. (2007). Mindfulness and emotion regulation: The development and initial validation of the cognitive and affective mindfulness scale-revised (CMS-R). *Journal of Psychopathology and Behavioral Assessment, 29*(3), 177-190.

Feldman, C., & Kuyken, W. (2011). Compassion in the landscape of suffering. *Contemporary Buddhism*, *12*(01), 143-155.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th Ed.). Sage Publications.

Fredrickson, B. L., Cohn, M. A., Coffey, K. A., Pek, J., & Finkel, S. M. (2008). Open hearts build lives: positive emotions, induced through loving-kindness meditation, build consequential personal resources. *Journal of personality and social psychology, 95*(5), 1045.

Fresco, D. M., Moore, M. T., van Dulmen, Manfred H. M., Segal, Z. V., Ma, S. H., Teasdale, J. D., & Williams, J. M. (2007). Initial psychometric properties of the experiences questionnaire: Validation of a self-report measure of decentering. *Behavior Therapy, 38*(3), 234-246.

Furr, M. (2011). *Scale construction and psychometrics for social and personality psychology*. SAGE Publications Ltd.

Galek, K., Flannelly, K. J., Vane, A., & Galek, R. M. (2005). Assessing a patient's spiritual needs: A comprehensive instrument. *Holistic Nursing Practice, 19*(2), 62-69.

Gilbert, P. (2005). Compassion and cruelty: A biopsychosocial approach. In: P. Gilbert

(Ed.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 9-74). London: Routledge.

Gilbert, P. (2009a). *The Compassionate Mind: A new approach to life's challenges*. London: Constable and Robinson.

Gilbert, P. (2009b). Introducing compassion-focused therapy. *Advances in psychiatric treatment, 15*(3), 199-208.

Gilbert, P. (2010). *The compassionate mind.* London: Constable & Robinson Ltd.

Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology, 53*(1), 6-41.

Gilbert, P., Catarino, F., Duarte, C., Matos, M., Kolts, R., Stubbs, J., ... & Basran, J. (2017). The development of compassionate engagement and action scales for self and others. *Journal of Compassionate Health Care, 4*(1), 4.

Gilbert, P., Clarke, M., Hempel, S., Miles, J. N. V., & Irons, C. (2004). Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. *British Journal of Clinical Psychology, 43*(1), 31-50.

Gilbert, P. & Irons, C. (2005). Focused therapies and compassionate mind training for shame and self-attacking. In: P. Gilbert (Ed.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 263-325). London: Routledge.

Gilbody, S., Littlewood, E., Hewitt, C., Brierley, G., Tharmanathan, P., Araya, R., ... & Kessler, D. (2015). Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. *Bmj, 351,* h5627.

Gillath, O., Shaver, P. R., & Mikulincer, M. (2005). An attachment-theoretical approach to compassion and altruism. In In P. Gilbert (Eds.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 122–147). London: Routledge.

Godfrin, K. A., & van Heeringen, C. (2010). The effects of mindfulness-based cognitive therapy on recurrence of depressive episodes, mental health and quality of life: A randomized controlled study. *Behaviour Research and Therapy, 48*(8), 738-746.

Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: an evolutionary analysis and empirical review. *Psychological bulletin, 136*(3), 351.

Goldberg, S. B., Tucker, R. P., Greene, P. A., Davidson, R. J., Wampold, B. E.,

Kearney, D. J., & Simpson, T. L. (2018). Mindfulness-based interventions for psychiatric disorders: A systematic review and meta-analysis. *Clinical psychology review, 59*, 52-60.

Goleman, D. (2003). *Destructive Emotions and How We Can Overcome Them: A Dialogue with the Dalai Lama*. London: Bloomsbury.

Grabovac, A. D., Lau, M. A., & Willett, B. R. (2011). Mechanisms of mindfulness: A Buddhist psychological model. *Mindfulness, 2*(3), 154-166.

Green, S. M., & Bieling, P. J. (2012). Expanding the scope of mindfulness-based cognitive therapy: Evidence for effectiveness in a heterogeneous psychiatric sample. *Cognitive and Behavioral Practice, 19*(1), 174-180.

Griffin, D. W., & Bartholomew, K. (1994). *Relationship Scales Questionnaire*. London, UK: Jessica Kingsley Publishers.

Grossman, P. (2008). On measuring mindfulness in psychosomatic and psychological research. *Journal of psychosomatic research, 64*(4), 405-408.

Grossman, P., Niemann, L., Schmidt, S., & Walach, H. (2004). Mindfulness-based stress reduction and health benefits: A meta-analysis. *Journal of psychosomatic research, 57*(1), 35-43.

Grossman, P., Tiefenthaler-Gilmer, U., Raysz, A., & Kesper, U. (2007). Mindfulness training as an intervention for fibromyalgia: Evidence of postintervention and 3-year follow-up benefits in well-being. *Psychotherapy and Psychosomatics, 76*(4), 226-233.

Grossman, P., & Van Dam, N. T. (2011). Mindfulness, by any other name…: trials and tribulations of sati in western psychology and science. *Contemporary Buddhism, 12*(01), 219-239.

Gu, J., Baer, R., Cavanagh, K., Kuyken, W., & Strauss, C. Developing and validating new self-report measures of compassion: Sussex-Oxford Compassion for Others Scale (SOCS-O) and Sussex-Oxford Compassion for the Self Scale (SOCS-S). *Manuscript under review.*

Gu, J., Cavanagh, K., Baer, R., & Strauss, C. (2017). An empirical examination of the factor structure of compassion. *PloS one, 12*(2), e0172471.

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical Psychology Review, 37*, 1-12.

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2016). Corrigendum to "How do Mindfulness-Based Cognitive Therapy and Mindfulness-Based Stress Reduction Improve Mental Health and Wellbeing? A Systematic Review and Meta-Analysis of Mediation Studies" [Clinical Psychology Review 37 (2015) 1-12]. *Clinical psychology review, 49,* 119.

Gu, J., Strauss, C., Crane, C., Barnhofer, T., Karl, A., Cavanagh, K., & Kuyken, W. (2016). Examining the factor structure of the 39-item and 15-item versions of the five facet mindfulness questionnaire before and after mindfulness-based cognitive therapy for people with recurrent depression. *Psychological Assessment, 28*(7), 791-802.

**Hacker, T. (2008). The relational compassion scale: development and validation of a new self rated scale for the assessment of self-other compassion (Doctoral dissertation, University of Glasgow).

Hassard, J., Teoh, K. R., Visockaite, G., Dewe, P., & Cox, T. (2018). The cost of work-related stress to society: A systematic review. *Journal of occupational health psychology, 23*(1), 1.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. Communication monographs, 76(4), 408-420.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology, 67*(3), 451–470.

Hayes, A. M., & Feldman, G. (2004). Clarifying the construct of mindfulness in the context of emotion regulation and the process of change in therapy. *Clinical Psychology: Science and Practice, 11*, 255–262.

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy, 44*(1), 1-25.

Hayes, S. C., & Wilson, K. G. (1994). Acceptance and commitment therapy: Altering the verbal support for experiential avoidance. *The Behavior Analyst, 17*(2), 289-303.

Health and Safety Executive. (2017). *Health and safety at work: Summary statistics for Great Britain 2017*. Retrieved from hse.gov.uk/statistics/overall/hssh1617.pdf

*Heeren, A., & Philippot, P. (2011). Changes in ruminative thinking mediate the clinical benefits of mindfulness: Preliminary findings. *Mindfulness, 2*(1), 8-13.

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British journal of clinical psychology, 44*(2), 227-239.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational research methods, 1*(1), 104-121.

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*, 169–183.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*(3), 307-316.

Hölzel, B.K., Lazar, S.W., Gard, T., Schuman-Olivier, Z., Vago, D.R., & Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspectives on Psychological Science, 6*, 537-559.

Hoogendoorn, W. E., van Poppel, ,M.N., Bongers, P. M., Koes, B. W., & Bouter, L. M. (2000). Systematic review of psychosocial factors at work and private life as risk factors for back pain. *Spine, 25*(16), 2114-2125.

Howells, A., Ivtzan, I., & Eiroa-Orosa, F. J. (2016). Putting the 'app' in happiness: a randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. *Journal of Happiness Studies, 17*(1), 163-185.

Hoyer, J., Gloster, A. T., & Herzberg, P. Y. (2009). Is worry different from rumination? yes, it is more predictive of psychopathology! *GMS Psycho-Social-Medicine, 6*, 1-9.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Hutcherson, C. A., Seppala, E. M., & Gross, J. J. (2008). Loving-kindness meditation increases social connectedness. *Emotion, 8*(5), 720.

**Hwang, J. Y., Plante, T., & Lackey, K. (2008). The development of the Santa Clara brief compassion scale: An abbreviation of Sprecher and Fehr's compassionate love scale. *Pastoral Psychology, 56*(4), 421-428.

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence

that structural equations models perform better than regressions. *Journal of Consumer Psychology, 17*(2), 139-153.

IBM. (2013). *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, NY: IBM Corp.

Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century.* Washington, DC: National Academies Press.

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials, 17*(1), 1-12.

Jazaieri, H., Jinpa, G. T., McGonigal, K., Rosenberg, E. L., Finkelstein, J., Simon-Thomas, E., ... & Goldin, P. R. (2013). Enhancing compassion: A randomized controlled trial of a compassion cultivation training program. *Journal of Happiness Studies, 14*(4), 1113-1126.

Jazaieri, H., McGonigal, K., Jinpa, T., Doty, J. R., Gross, J. J., & Goldin, P. R. (2014). A randomized controlled trial of compassion cultivation training: Effects on mindfulness, affect, and emotion regulation. *Motivation and Emotion, 38*(1), 23-35.

Jimenez, S. (2008). The role of self-acceptance, negative mood regulation, and ruminative brooding on mindfulness and depressive symptoms: a longitudinal, randomized controlled trial of mindfulness meditation vs. relaxation training. Doctoral Dissertations, AAI3325617.

Jinpa, G. T. (2010). Compassion cultivation training (CCT): Instructor's manual (unpublished manual). The Center for Compassion and Altruism Research and Education, Stanford University, CA.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In: L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research,* (pp. 102-138). New York: Guildford Press.

Joinson, C. (1992). Coping with compassion fatigue. *Nursing, 22*(4), 116-118.

Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General Hospital Psychiatry, 4*(1), 33-47.

Kabat-Zinn, J. (1994). *Wherever you go, there you are: Mindfulness meditation in*

everyday life. New York: Hyperion Books.

Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: past, present, and future. *Clinical psychology: Science and practice, 10*(2), 144-156.

Kanov, J. M., Maitlis, S., Worline, M. C., Dutton, J. E., Frost, P. J., & Lilius, J. M. (2004). Compassion in organizational life. *American Behavioral Scientist, 47*(6), 808-827.

Kanske, P., Böckler, A., Trautwein, F. M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and Theory of Mind. *NeuroImage, 122*, 6-19.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27.

Keltner, D. (2009). *Born to be good: The science of a meaningful life*. New York: Norton & Company.

Kemeny, M. E., Foltz, C., Cavanagh, J. F., Cullen, M., Giese-Davis, J., Jennings, P., ... & Ekman, P. (2012). Contemplative/emotion training reduces negative emotional behavior and promotes prosocial responses. *Emotion, 12*(2), 338.

*Keng, S., Smoski, M. J., Robins, C. J., Ekblad, A. G., & Brantley, J. G. (2012). Mechanisms of change in mindfulness-based stress reduction: Self-compassion and mindfulness as mediators of intervention outcomes. *Journal of Cognitive Psychotherapy, 26*(3), 270-280.

Khoury, B., Sharma, M., Rush, S. E., & Fournier, C. (2015). Mindfulness-based stress reduction for healthy individuals: a meta-analysis. *Journal of psychosomatic research, 78*(6), 519-528.

Kiken, L. G., Garland, E. L., Bluth, K., Palsson, O. S., & Gaylord, S. A. (2015). From a state to a trait: trajectories of state mindfulness in meditation during intervention predict changes in trait mindfulness. *Personality and Individual differences, 81*, 41-46.

Kirby, J. N. (2017). Compassion interventions: the programmes, the evidence, and implications for research and practice. *Psychology and Psychotherapy: Theory, Research and Practice, 90*(3), 432-455.

Kirby, J. N., Tellegen, C. L., & Steindl, S. R. (2017). A meta-analysis of compassion-based interventions: Current state of knowledge and future directions. *Behavior therapy, 48*(6), 778-792.

Klimecki, O. M., Leiberg, S., Ricard, M., & Singer, T. (2013). Differential pattern of

functional brain plasticity after compassion and empathy training. *Social cognitive and affective neuroscience, 9*(6), 873-879.

Kline, P. (1999). *The handbook of psychological testing* (2nd Ed.). London, UK: Routledge.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd Ed.). New York: Guilford Press.

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th Ed.). New York: Guilford Press.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877-883.

Kuyken, W., Byford, S., Taylor, R. S., Watkins, E., Holden, E., White, K., Barrett, B., Byng, R., Evans, A., Mullan, E., & Teasdale, J. D. (2008). Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *Journal of Consulting and Clinical Psychology, 76*, 966–978.

Kuyken, W., Hayes, R., Barrett, B., Byng, R., Dalgleish, T., Kessler, D., ... & Causley, A. (2015). Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): a randomised controlled trial. *The Lancet, 386*(9988), 63-73.

Kuyken, W., Warren, F. C., Taylor, R. S., Whalley, B., Crane, C., Bondolfi, G., ... & Segal, Z. (2016). Efficacy of mindfulness-based cognitive therapy in prevention of depressive relapse: An individual patient data meta-analysis from randomized trials. *JAMA psychiatry*, *73*(6), 565-574.

*Kuyken, W., Watkins, E., Holden, E., White, K., Taylor, R. S., Byford, S., Evans, A., Radford, S., Teasdale, J. D., & Dalgleish, T. (2010). How does mindfulness-based cognitive therapy work? *Behaviour Research and Therapy, 48*(11), 1105-1112.

Labbé, E., Schmidt, N., Babin, J., & Pharr, M. (2007). Coping with stress: The effectiveness of different types of music. *Applied Psychophysiology and Biofeedback, 32*(3-4), 163-168.

*Labelle, L. E. (2012). How does mindfulness-based stress reduction (MBSR) improve psychological functioning in cancer patients? (Order No. NR87924, University of Calgary (Canada)). ProQuest Dissertations and Theses, 272.

*Labelle, L. E., Campbell, T. S., & Carlson, L. E. (2010). Mindfulness-based stress reduction in oncology: Evaluating mindfulness and rumination as mediators of change in depressive symptoms. *Mindfulness, 1*(1), 28-40.

Lau, M. A., Bishop, S. R., Segal, Z. V., Buis, T., Anderson, N. D., Carlson, L., ... & Devins, G. (2006). The Toronto mindfulness scale: Development and validation. *Journal of clinical psychology, 62*(12), 1445-1467.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford: Oxford University Press.

Ledesma, D., & Kumano, H. (2009). Mindfulness-based stress reduction and cancer: A meta-analysis. *Psycho-Oncology, 18*(6), 571-579.

Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [Software]. Available from quantpsy.org/corrtest/corrtest2.htm.

Lee, R. M., & Robbins, S. B. (1995). Measuring belongingness: The Social Connectedness and the Social Assurance scales. *Journal of Counseling Psychology, 42*(2), 232.

Leiberg, S., Klimecki, O., & Singer, T. (2011). Short-term compassion training increases prosocial behavior in a newly developed prosocial game. *PloS one, 6*(3), e17798.

*Lengacher, C. A., Shelton, M. M., Reich, R. R., Barta, M. K., Johnson-Mallard, V., Moscoso, M. S., . . . Kip, K. E. (2014). Mindfulness based stress reduction (MBSR(BC)) in breast cancer: Evaluating fear of recurrence (FOR) as a mediator of psychological and physical symptoms in a randomized control trial (RCT). *Journal of Behavioral Medicine, 37*(2), 185-195.

Levine, T. R. (2005). Confirmatory factor analysis and scale validation in communication research. *Communication Research Reports, 22*(4), 335-338.

Levy, P. (1967). The correction for spurious correlation in the evaluation of short-form tests. *Journal of Clinical Psychology, 23*, 84-86.

Lim, D., Condon, P., & DeSteno, D. (2015). Mindfulness and compassion: an examination of mechanism and scalability. *PloS one, 10*(2), e0118221.

Linehan, M. M. (1993). *Cognitive–behavioral therapy of borderline personality disorder*. New York: Guilford Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or

not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151–173.

Lloyd, J., Bond, F. W., & Flaxman, P. E. (2013). The value of psychological flexibility: Examining psychological mechanisms underpinning a cognitive behavioural therapy intervention for burnout. *Work & Stress, 27*(2), 181-199.

Lo, C. S. L., Ho, S. M. Y., & Hollon, S. D. (2008). The effects of rumination and negative cognitive styles on depression: A mediation analysis. *Behaviour Research and Therapy, 46*(4), 487-495.

López, A., Sanderman, R., Ranchor, A. V., & Schroevers, M. J. (2018). Compassion for others and self-compassion: Levels, correlates, and relationship with psychological well-being. *Mindfulness, 9*(1), 325-331.

López, A., Sanderman, R., Smink, A., Zhang, Y., van Sonderen, E., Ranchor, A., & Schroevers, M. J. (2015). A reconsideration of the Self-Compassion Scale's total score: Self-compassion versus self-criticism. *PloS One, 10*(7), e0132940.

**Lown, B. A., Muncer, S. J., & Chadwick, R. (2015). Can compassionate healthcare be measured? The Schwartz Center Compassionate Care Scale™. *Patient education and counseling, 98*(8), 1005-1010.

Lown, B. A., Rosen, J., & Marttila, J. (2011). An agenda for improving compassionate care: a survey shows about half of patients say such care is missing. *Health Affairs, 30*(9), 1772-1778.

Lubans, D. R., Foster, C., & Biddle, S. J. H. (2008). A review of mediators of behavior in interventions to promote physical activity among children and adolescents. *Preventive Medicine, 47*(5), 463-470.

Ludwig, D. S., & Kabat-Zinn, J. (2008). Mindfulness in medicine. *Jama, 300*(11), 1350-1352.

Ly, K. H., Trüschel, A., Jarl, L., Magnusson, S., Windahl, T., Johansson, R., ... & Andersson, G. (2014). Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. *BMJ open, 4*(1), e003440.

MacBeth, A. & Gumley, A. (2012). Exploring compassion: a meta-analysis of the association between self-compassion and psychopathology. *Clinical Psychology Review, 32*, 545-552.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods, 4*(1), 84.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*(1), 83-104.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*(1), 99-128.

Mardia, K. V. (1985). Mardia's test of multinormality. In: S. Kotz & N. L. Johnson (Eds.), *Encyclopaedia of statistical sciences* (Vol. 5, pp. 217–221). New York: Wiley.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling, 11*(3), 320-341.

**Martins, D., Nicholas, N. A., Shaheen, M., Jones, L., & Norris, K. (2013). The development and evaluation of a compassion scale. *Journal of health care for the poor and underserved, 24*(3), 1235.

Mascaro, J. S., Rilling, J. K., Tenzin Negi, L., & Raison, C. L. (2012). Compassion meditation enhances empathic accuracy and related neural activity. *Social cognitive and affective neuroscience, 8*(1), 48-55.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1981). *Maslach Burnout Inventory: MBI*. Palo Alto, CA: Consulting psychologists press.

Matsunaga, M. (2010). How to Factor-Analyze Your Data Right: Do's, Don'ts, and How-To's. *International journal of psychological research, 3*(1).

McCrone, P. R., Dhanasiri, S., Patel, A., Knapp, M., & Lawton-Smith, S. (2008). *Paying the price: the cost of mental health care in England to 2026*. London: King's Fund.

*McManus, F., Surawy, C., Muse, K., Vazquez-Montes, M., & Williams, J. M. (2012). A randomized clinical trial of mindfulness-based cognitive therapy versus unrestricted services for health anxiety (hypochondriasis). *Journal of Consulting and Clinical Psychology, 80*(5), 817-828.

Meade, A. W., & Bauer, D. J. (2007). Power and precision in confirmatory factor analytic tests of measurement invariance. *Structural Equation Modeling, 14*(4), 611-635.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of*

Personality, 40(4), 525-543.

Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological bulletin, 111*(1), 172-175.

Mental Health Foundation. (2018). *Stress: Are we coping?* London: Mental Health Foundation.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy, 28*(6), 487-495.

Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., ... & Kaplan, R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and psychosomatics, 78*(5), 275-284.

Mongrain, M., Chin, J. M., & Shapira, L. B. (2011). Practicing compassion increases happiness and self-esteem. *Journal of Happiness Studies, 12*(6), 963-981.

Mongrain, M., Komeylian, Z., & Barnhart, R. (2016). Happiness vs. mindfulness exercises for individuals vulnerable to depression. *The Journal of Positive Psychology, 11*(4), 366–377.

Moses, J. (2002). *Oneness: Great principles shared by all religions.* Toronto: Ballantine Books.

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's Guide* (7th Ed.). Los Angeles, CA: Muthén & Muthén.

Najjar, N., Davis, L. W., Beck-Coon, K., & Carney Doebbeling, C. (2009). Compassion fatigue: a review of the research to date and relevance to cancer-care providers. *Journal of Health Psychology, 14*, 267–77.

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews of the Cambridge Philosophical Society, 82*(4), 591-605.

Neff, K. D. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity, 2*(2), 85-101.

**Neff, K. D. (2003b). The development and validation of a scale to measure self-compassion. *Self and Identity, 2*(3), 223–250.

Neff, K. D. (2015). The Self-Compassion Scale is a Valid and Theoretically Coherent Measure of Self-Compassion. *Mindfulness*, 1-11.

Neff, K. D., & Germer, C. K. (2013). A pilot study and randomized controlled trial of

the mindful self-compassion program. *Journal of clinical psychology, 69*(1), 28-44.

Neff, K. D., & Pommier, E. (2013). The relationship between self-compassion and other-focused concern among college undergraduates, community adults, and practicing meditators. *Self and Identity, 12*(2), 160-176.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling, 8*(3), 353-377.

NHS England. (2017). NHS staff health & wellbeing. Retrieved from www.england.nhs.uk/wp-content/uploads/2017/10/staff-health-wellbeing-cquin-2017-19-implementation-support.pdf

Niles, B. L., Klunk-Gillis, J., Ryngala, D. J., Silberbogen, A. K., Paysnick, A., & Wolf, E. J. (2012). Comparing mindfulness and psychoeducation treatments for combat-related PTSD using a telehealth approach. *Psychological Trauma: Theory, Research, Practice, and Policy, 4*(5), 538.

Norko, M. A. (2005). Commentary: Compassion at the core of forensic ethics. *Journal of the American Academy of Psychiatry and the Law, 33*, 386–389.

*Nyklicek, I., Beugen, S., & Denollet, J. (2013). Effects of mindfulness-based stress reduction on distressed (type D) personality traits: A randomized controlled trial. *Journal of Behavioral Medicine, 36*(4), 361-370.

*Nyklicek, I., & Kujipers, K. F. (2008). Effects of Mindfulness-Based Stress Reduction Intervention on Psychological Wellbeing and Quality of Life: Is Increased Mindfulness Indeed the Mechanism? *Annals of Behavioral Medicine, 35*(3), 331-340.

O'Connor, R. C., & Noyce, R. (2008). Personality and cognitive processes: Self-criticism and different types of rumination as predictors of suicidal ideation. *Behaviour Research and Therapy, 46*(3), 392-401.

Pace, T. W., Negi, L. T., Adame, D. D., Cole, S. P., Sivilli, T. I., Brown, T. D., ... & Raison, C. L. (2009). Effect of compassion meditation on neuroendocrine, innate immune and behavioral responses to psychosocial stress. *Psychoneuroendocrinology, 34*(1), 87-98.

Papageorgiou, C., & Wells, A. (1999). Process and metacognitive dimensions of depressive and anxious thoughts and relationships with emotional intensity. *Clinical Psychology and Psychotherapy, 6*, 156-162.

Pearson, A. (2006). Powerful caring. *Nursing Standard (Royal College of Nursing (Great Britain): 1987), 20*(48), 20-22.

Penner, L. A., Fritzsche, B. A., Craiger, J. P., & Freifeld, T. R. (1995). Measuring the prosocial personality. In J. Butcher, & C. D. Spielberger (Eds.), *Advances in personality assessment*, *Vol. 10.* Hillsdale, NJ: LEA.

Piet, J., & Hougaard, E. (2011). The effect of Mindfulness-Based Cognitive Therapy for prevention of relapse in recurrent major depressive disorder: A systematic review and meta-analysis. *Clinical Psychology Review, 31*, 1032–1040.

Plante, T. G., & Boccaccini, M. T. (1997). The Santa Clara strength of religious faith questionnaire. *Pastoral Psychology, 45*(5), 375–387.

**Pommier, E. A. (2010). The compassion scale. The University of Texas at Austin.

Porges, S. W. (2009). The polyvagal theory: new insights into adaptive reactions of the autonomic nervous system. *Cleveland Clinic journal of medicine, 76*(Suppl 2), S86.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*(4), 717-731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879-891.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*(1), 185-227.

Quaglia, J. T., Braun, S. E., Freeman, S. P., McDaniel, M. A., & Brown, K. W. (2016). Meta-analytic evidence for effects of mindfulness training on dimensions of self-reported dispositional mindfulness. *Psychological assessment, 28*(7), 803.

R Development Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

*Raes, F., Dewulf, D., Van Heeringen, C., & Williams, J. M. (2009). Mindfulness and reduced cognitive reactivity to sad mood: Evidence from a correlational study and a non-randomized waiting list controlled study. *Behaviour Research and Therapy, 47*(7), 623-627.

Raes, F., Hermans, D., Williams, J. M. G., Beyers, W. Brunfaut, El., & Eelen, P.

(2006). Reduced Autobiographical Memory Specificity and Rumination in Predicting the Course of Depression. *Journal of Abnormal Psychology, 115*(4), 699-704.

**Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the self-compassion scale. *Clinical psychology & psychotherapy, 18*(3), 250-255.

Rahula, W. (1959). *What the Buddha taught*. New York: Grove Press.

Rapgay, L., & Bystrisky, A. (2009). Classical mindfulness. *Annals of the New York Academy of Sciences, 1172*(1), 148-162.

Redelmeier, D. A., Molin, J. P., & Tibshirani, R. J. (1995). A randomised trial of compassionate care for the homeless in an emergency department. *The Lancet, 345*(8958), 1131–1134.

Robins, C. J., Keng, S. L., Ekblad, A. G., & Brantley, J. G. (2012). Effects of mindfulness-based stress reduction on emotional experience and expression: A randomized controlled trial. *Journal of clinical psychology, 68*(1), 117-131.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton UP.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass, 5*(6), 359-371.

Rycroft-Malone, J., Gradinger, F., Griffiths, H. O., Crane, R., Gibson, A., Mercer, S., ... & Kuyken, W. (2017). Accessibility and implementation in the UK NHS services of an effective depression relapse prevention programme: learning from mindfulness-based cognitive therapy through a mixed-methods study. *Health Services and Delivery Research, 5*(14), 1-190.

Safran, J. D., & Segal, Z. V. (1990). *Interpersonal process in cognitive therapy*. New York: Basic Books.

Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. *Emotion, disclosure, and health, 125*, 154.

Sanghavi, D. M. (2006). What makes for a compassionate patient-caregiver relationship? *Joint Commission journal on quality and patient safety, 32*(5), 283-292.

Sauer, S., Walach, H., Schmidt, S., Hinterberger, T., Lynch, S., Büssing, A., & Kohls,

N. (2013). Assessment of mindfulness: Review on state of the art. *Mindfulness, 4*(1), 3-17.

Scher, C. D., Ingram, R. E., & Segal, Z. V. (2005). Cognitive reactivity and vulnerability: Empirical evaluation of construct activation and cognitive diatheses in unipolar depression. *Clinical Psychology Review, 25*(4), 487-510.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*, 23–74.

Schieman, S., & Van Gundy, K. (2000). The personal and social links between age and self-reported empathy. *Social Psychology Quarterly, 63*, 152–174.

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *PLoS Medicine, 7*(3), 1.

Segal, Z. V., Williams, J. M., & Teasdale, J. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse*. London, UK: Guilford Press.

Segal, Z. V., Williams, J. M., & Teasdale, J. (2013). *Mindfulness-based cognitive therapy for depression* (2nd Ed.). London, UK: Guilford Press.

Shadish, W. R. (1996). Meta-analysis and the exploration of causal mediating processes: A primer of examples, methods, and issues. *Psychological Methods, 1*(1), 47-65.

*Shahar, B., Britton, W. B., Sbarra, D. A., Figueredo, A. J., & Bootzin, R. R. (2010). Mechanisms of change in mindfulness-based cognitive therapy for depression: Preliminary evidence from a randomized controlled trial. *International Journal of Cognitive Therapy, 3*(4), 402-418.

Shapiro, S. L., Astin, J. A., Bishop, S. R., & Cordova, M. (2005). Mindfulness-based stress reduction for health care professionals: results from a randomized trial. *International Journal of Stress Management, 12*(2), 164.

Shapiro, S. L., Carlson, L. E., Astin, J. A., & Freedman, B. (2006). Mechanisms of mindfulness. *Journal of Clinical Psychology, 62*(3), 373-386.

Shapiro, S. L., Schwartz, G. E., & Bonner, G. (1998). Effects of mindfulness-based stress reduction on medical and premedical students. *Journal of behavioral medicine, 21*(6), 581-599.

Siegel, R. D., Germer, C. K., & Olendzki, A. (2009). Mindfulness: What is it? Where

did it come from?. In *Clinical handbook of mindfulness* (pp. 17-35). New York: Springer.

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology, 24*(18), R875-R878.

Singh, A. S., Mulder, C., Twisk, J. W. R., van Mechelen, W., & Chinapaw, M. J. M. (2008). Tracking of childhood overweight into adulthood: A systematic review of the literature. *Obesity Reviews, 9*(5), 474-488.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*(1), 102-111.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology, 13*, 290-312.

Spielberger, C. D., Gorsuch, R.L., and Lushene. R.E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spijkerman, M. P. J., Pots, W. T. M., & Bohlmeijer, E. T. (2016). Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. *Clinical psychology review, 45*, 102-114.

**Sprecher, S., & Fehr, B. (2005). Compassionate love for close others and humanity. *Journal of Social and Personal Relationships, 22*(5), 629-651.

Stankovic, D. (2015). Mindfulness meditation training for tennis players. Doctoral dissertation, Boston University.

Stanton, A. L., Kirk, S. B., Cameron, C. L., & Danoff-Burg, S. (2000). Coping through emotional approach: Scale construction and validation. *Journal of Personality and Social Psychology, 78*, 1150–1169.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological bulletin, 87*(2), 245-251.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173-180.

Stewart, L. A., & Parmar, M. K. B. (1993). Meta-analysis of the literature or of individual patient data: is there a difference? *The Lancet, 341*(8842), 418–422.

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S.

(2009). Internal construct validity of the Warwick-Edinburgh mental well-being scale (WEMWBS): a Rasch analysis using data from the Scottish health education population survey. *Health and quality of life outcomes, 7*(1), 15.

Strauss, C., Cavanagh, K., Oliver, A., Pettman, D. (2014). Mindfulness-Based Interventions for People Diagnosed with a Current Episode of an Anxiety or Depressive Disorder: A Meta-Analysis of Randomised Controlled Trials. *PLoS ONE, 9*(4), e96110.

Strauss, C., Gu, J., Kuyken, W., Cavanagh, K., & Baer, R. A meta-analysis of randomised controlled trials of the effects of mindfulness-based interventions in comparison to active control conditions on depressive symptoms at post-intervention and follow-up. *Manuscript in preparation.*

Strauss, C., Gu, J., Pitman, N., Chapman, C., Kuyken, W., & Whittington, A. A randomised controlled trial of mindfulness-based cognitive therapy for life to improve healthcare staff stress. *Manuscript in preparation.*

Strauss, C., Taylor, B. L., Gu, J., Kuyken, W., Baer, R., Jones, F., & Cavanagh, K. (2016). What is compassion and how can we measure it? A review of definitions and measures. *Clinical psychology review, 47*, 15-27.

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition, 131*(1), 159-171.

Tanay, G., & Bernstein, A. (2013). State Mindfulness Scale (SMS): development and initial validation. *Psychological assessment, 25*(4), 1286.

Teasdale, J. D. (1988). Cognitive vulnerability to persistent depression. *Cognition and Emotion, 2*, 247–274.

Teasdale, J. D. (1999). Metacognition, mindfulness and the modification of mood disorders. *Clinical Psychology and Psychotherapy, 6*, 146-155.

Teasdale, J. D., & Chaskalson, M. (2011). How does mindfulness transform suffering? I: The nature and origins of dukkha. *Contemporary Buddhism, 12*(1), 89-102.

Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology, 8,* 615–623.

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L.,

Dekker, J.,...de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*, 34–42.

Thurman, R. (1997). *Essential Tibetan Buddhism*. New York: HarperCollins Publishers Inc.

Tirch, D. D. (2010). Mindfulness as a context for the cultivation of compassion. *International Journal of Cognitive Therapy, 3*(2), 113-123.

Trapnell, P. D., & Campbell, J. D. (1999). Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. *Journal of Personality and Social Psychology, 76*(2), 284-304.

Trautwein, F. M., Kanske, P., Böckler-Raettig, A., & Singer, T. (2017). Differential benefits of mental training types for attention, compassion, and theory of mind.

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination reconsidered: A psychometric analysis. *Cognitive Therapy and Research, 27*(3), 247-259.

Turner, E. H. (2013). Publication bias, with a focus on psychiatry: Causes and solutions. *CNS Drugs, 27*(6), 457-468.

Underwood, L. G., & Teresi, J. A. (2002). The daily spiritual experience scale: development, theoretical description, reliability, exploratory factor analysis, and preliminary construct validity using health-related data. *Annals of Behavioral Medicine, 24*(1), 22–33.

Vago, D. R., & Silbersweig, D. A. (2012). Self-awareness, self-regulation, and self-transcendence (S-ART): A framework for understanding the neurobiological mechanisms of mindfulness. *Frontiers in Human Neuroscience, 6*, 1-30.

Valk, S. L., Bernhardt, B. C., Trautwein, F. M., Böckler, A., Kanske, P., Guizard, N., ... & Singer, T. (2017). Structural plasticity of the social brain: Differential change after socio-affective and cognitive mental training. *Science Advances, 3*(10), e1700489.

*van Aalderen, J. R., Donders, A. R. T., Giommi, F., Spinhoven, P., Barendregt, H. P., & Speckens, A. E. M. (2012). The efficacy of mindfulness-based cognitive therapy in recurrent depressed patients with and without a current depressive episode: A randomized controlled trial. *Psychological Medicine, 42*(5), 989-1001.

van den Brink, E. & Koster, F. (2015). *Mindfulness-based Compassionate Living: A*

New Training Programme to Deepen Mindfulness with Heartfulness. Routledge.

van Stralen, M. M., Yildirim, M., Te Velde, S. J., Brug, J., van Mechelen, W., & Chinapaw, M. J. M. (2011). What works in school-based energy balance behaviour interventions and what does not? A systematic review of mediating mechanisms. *International Journal of Obesity, 35*, 1251-1265.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate behavioral research, 25*(1), 1-28.

Verplanken, B., & Fisher, N. (2013). Habitual worrying and benefits of mindfulness. *Mindfulness, 5,* 566–573.

*Vøllestad, J., Sivertsen, B., & Nielsen, G. H. (2011). Mindfulness-based stress reduction for patients with anxiety disorders: Evaluation in a randomized controlled trial. *Behaviour Research and Therapy, 49*(4), 281-288.

Wahbeh, H., Goodrich, E., & Oken, B. S. (2016). Internet mindfulness meditation for cognition and mood in older adults: a pilot study. *Alternative therapies in health and medicine, 22*(2), 44.

Walach, H., Buchheld, N., Buttenmuller, V., Kleinknecht, N., & Schmidt, S. (2006). Measuring mindfulness: The Freiburg Mindfulness Inventory (FMI). *Personality and Individual Differences, 40*, 1543–1555.

Wallace, B. A. (2001). Intersubjectivity in Indo-Tibetan Buddhism. Journal of *Consciousness Studies, 8*(5-6), 209-230.

Wang, S. (2005). A conceptual framework for integrating research related to the physiology of compassion and the wisdom of Buddhist teachings. In P. Gilbert (Eds.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 75–120). London: Routledge.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin, 134*(2), 163-206.

Watkins, E., Moulds, M., & Mackintosh, B. (2005). Comparisons between rumination and worry in a non-clinical population. *Behaviour Research and Therapy, 43*(12), 1577-1585.

Watkins, E., & Teasdale, J. D. (2001). Rumination and overgeneral memory in depression: Effects of self-focus and analytic thinking. *Journal of Abnormal Psychology, 110*, 353–357.

Watkins, E., & Teasdale, J. D. (2004). Adaptive and maladaptive self-focus in depression. *Journal of Affective Disorders, 82*, 1–8.

WHOQOL Group. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine, 28*, 551-558.

Wicksell, R. K., Olsson, G. L., & Hayes, S. C. (2010). Psychological flexibility as a mediator of improvement in acceptance and commitment therapy for patients with chronic pain following whiplash. *European Journal of Pain, 14*(10), e1-e11.

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10-18.

Williams, J. M. G., Crane, C., Barnhofer, T., Brennan, K., Duggan, D. S., Fennell, M. J. V., . . . Russell, I. T. (2014). Mindfulness-based cognitive therapy for preventing relapse in recurrent depression: A randomized dismantling trial. *Journal of Consulting and Clinical Psychology, 82*(2), 275-286.

Williams, M. J., Dalgleish, T., Karl, A., & Kuyken, W. (2014). Examining the factor structures of the five facet mindfulness questionnaire and the self-compassion scale. *Psychological Assessment, 26*(2), 407-418.

Williams, J. M., Teasdale, J. D., Segal, Z. V., & Soulsby, J. (2000). Mindfulness-based cognitive therapy reduces overgeneral autobiographical memory in formerly depressed patients. *Journal of Abnormal Psychology, 109*(1), 150-155.

Wispe, L. (1991). *The psychology of sympathy*. New York: Plenum.

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. Geneva: World Health Organization Document Production Services.

Wu, Y., Larrabee, J. H., & Putman, H. P. (2006). Caring behaviors inventory: A reduction of the 42-item instrument. *Nursing Research, 55*(1), 18-25.

Appendices

Appendix A: Supplementary Materials for Chapter 1

Table 1.
Definitions and known empirical support for theorised mechanisms of mindfulness and mindfulness-based interventions

Theorised mechanism	Definition	Empirical support[a]
Mindfulness	"the awareness that emerges through paying attention on purpose, in the present moment, and non-judgmentally to the unfolding experience moment by moment" (Kabat-Zinn, 2003, p. 145)	Mindfulness has been found to mediate the effects of MBCT/MBSR in a number of RCTs (e.g., Batink et al., 2013, Bränström et al., 2010, Kuyken et al., 2010).
Reperceiving (decentering, metacognitive insight)	A shift in perspective whereby one is able to experience mental content as events in awareness in the moment that they occur and disengage and disidentify from them.	None known
Repetitive negative thinking (rumination, worry)	A style of thinking about one's problems or negative experiences that is repetitive, at least partly intrusive, and difficult to disengage from (Ehring et al., 2011).	Rumination and worry have been found to medicate the effects of MBCT/MBSR in at least 5 RCTs (e.g., Batink et al., 2013, van Aalderen et al., 2012).
Reactivity	The extent to which a mild state of distress coupled with stress reactivates negative thinking and emotional patterns.	Emotional reactivity has been found to mediate the effects of MBCT (Britton et al., 2010).
Self-regulation	A process of maintaining stability of functioning and adaptability to change.	None known
Attention regulation	The ability to focus and sustain attention whilst disregarding distractions.	None known
Emotion regulation	Strategies for regulating emotional responses, e.g., reappraisal (changing interpretations of an experience) and extinction (exposing oneself to experiences).	None known
Psychological flexibility	Thinking, feeling, and behaving in ways that are more flexible and adaptive rather than automatic, reactive, and conditioned.	Psychological flexibility has been found to mediate the effects of MBSR (Labelle, 2012)
Values clarification	Greater recognition of what is meaningful to us so that we can reflectively make decisions that are congruent with our values, interests, and needs.	None known
Exposure	Desensitisation and understanding that emotions, thoughts, and bodily sensations are not as overwhelming as expected.	None known
Body awareness	The ability to recognise subtle bodily sensations.	None known
Self-awareness	Meta-awareness of the self.	None known
Self-transcendence	Developing a positive relationship between self and other that transcends self-focus and increases prosocial qualities.	None known
Compassion	Recognising suffering, understanding the universality of human suffering, feeling for the	Compassion for the self has been found to mediate

	person suffering, tolerating uncomfortable feelings, and motivation to act/acting to alleviate suffering.	the effects of MBCT (Kuyken et al., 2010).
Acceptance and non-attachment	Fully experiencing events as they are, without defence, avoidance, control, or investment of wellbeing in attaining goals or changing circumstances.	None known
Relaxation	The state of increased calmness and being free from tension in body and mind.	None known
Ethical practices	Code of ethics (e.g., not intentionally killing, stealing, lying, and using intoxicants) to reduce mental proliferation.	None known
Mind-body functioning	Health benefits gained through enhancing immunological resistance, relaxation, pain tolerance, and other physical processes.	None known

[a] Support for a theorised mechanism from controlled studies of MBCT or MBSR which conducted mediation analysis.

Appendix B: Supplementary Materials for Chapter 2

Table 1.
Quality checklist and scores assigned for RCTs

Quality criteria	Batink *et al.* (2013)	Bieling *et al.* (2012)	Bränström *et al.* (2010)	Britton *et al.* (2012)	Crane *et al.* (2012)	Keng *et al.* (2012)	Kuyken *et al.* (2010)	Lengacher *et al.* (2014)	McManus *et al.* (2012)	Nyklícek & Kuijpers (2008)	Nyklícek *et al.* (2013)	Shahar *et al.* (2010)	van Aalderen *et al.* (2012)	Vøllestad *et al.* (2011)
1) Did the study cite a theoretical framework?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2) Were the psychometric characteristics of the mediator(s) tested and were they within acceptable ranges (e.g., Cronbach's alpha/test-retest reliability > .7)?	No	Yes	Yes	Yes	No	No	No	No	Yes	No	No	No	No	Yes
3) Were the psychometric characteristics of the outcome(s) tested and were they within acceptable ranges (e.g., Cronbach's alpha/test-retest reliability > .7)?	No	No	Yes	Yes	Yes	No	No	No	Yes	No	No	Yes	No	Yes
4) Did the study report a power calculation and was the study adequately	Yes	No	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	No	No

powered to detect mediation?														
5) Did the study have an active control group?	No	Yes	No	No	No	No	Yes	No	No	No	No	No	No	No
6) Was the study described as randomised?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7) Was the method used to generate the sequence of randomisation described and appropriate (e.g., table of random numbers, computer-generated)?	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
8) Were participants or experimenters blind to treatment assignment?	No	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
9) Was information about participant flow provided (numbers of participants assigned to each condition, analysed, and dropped out)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

10) Did the study report the experience/qualifications of MBI facilitator(s)?	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No
11) Was change in the mediator(s) used in mediation analysis measured before the outcome?	No	No	No	No	No	No	Yes	No	No	No	No	No	No	No
12) Was change in the mediator(s) used in mediation analysis measured during treatment?	No	No	No	No	No	No	No	No	No	No	No	No	No	No
13) Did the study report the proportion of participants in the intervention condition who received an adequate dose of the MBI (> 4 out of 8 sessions)?	No	No	Yes	Yes	No	Yes	Yes	No	No	Yes	Yes	No	Yes	Yes
14) Was mediation analysis carried out using only the participants who received an adequate dose of the MBI?	No	No	No	Yes	No	Yes	Yes	No	No	No	No	No	Yes	No

15) Did post-intervention outcome control for baseline scores?	No	Yes	Yes	Yes	No	No	Yes	No	No	No	Yes	No	Yes	Yes
16) Was the most appropriate method of mediation analysis used, given the study design and recommendations?	No	Yes	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Quality score (/16)	6	10	10	13	5	7	12	6	10	7	9	9	10	8

Note. MBI = mindfulness-based intervention; RCT = randomised controlled trial.

Table 2.
Quality checklist and scores assigned for quasi-experimental studies

Quality criteria	Bergen-Cico & Cheon (2013)	Heeren & Philippot (2011)	Labelle (2012)	Labelle *et al.* (2010)	Raes *et al.* (2009)
1) Did the study cite a theoretical framework?	Yes	Yes	Yes	Yes	Yes
2) Were the psychometric characteristics of the mediator(s) tested and were they within acceptable ranges (e.g. Cronbach's alpha/test-retest reliability > .7)?	Yes	No	Yes	No	No
3) Were the psychometric characteristics of the outcome(s) tested and were they within acceptable ranges (e.g. Cronbach's alpha/test-retest reliability > .7)?	Yes	No	Yes	No	No
4) Did the study report a power calculation and was the study adequately powered to detect mediation?	No	No	Yes	No	No
5) Did the study have an active control group?	Yes	No	No	No	No
6) Was the study described as randomised?	No	No	No	No	No
7) Was the method used to generate the sequence of randomisation described and appropriate (e.g., table of random numbers, computer-generated)?	No	No	No	No	No
8) Were participants or experimenters blind to treatment assignment?	No	No	No	No	No
9) Was information about participant flow provided (numbers of participants assigned to each condition, analysed, and dropped out)?	Yes	Yes	Yes	Yes	No
10) Did the study report the experience/qualifications of the MBI facilitator(s)?	No	Yes	No	No	No
11) Was change in the mediator(s) used in mediation analysis measured before the outcome?	Yes	No	No	No	No
12) Was change in the mediator(s) used in mediation analysis measured during treatment?	No	No	No	No	No
13) Did the study report the proportion of participants in the intervention condition who received an adequate dose of the MBI (> 4 out of 8 sessions)?	No	Yes	Yes	Yes	No
14) Was mediation analysis carried out using only the participants who received an adequate dose of the MBI?	No	Yes	Yes	No	No
15) Did post-intervention outcome control for baseline scores?	Yes	Yes	No	Yes	Yes
16) Was the most appropriate method of mediation analysis used, given the study design and recommendations?	No	No	Yes	Yes	Yes
Quality score (/16)	7	6	8	5	3

Note. MBI = mindfulness-based intervention.

Table 3.
Descriptions of included randomised controlled trials (RCTs)

Study	Sample (*N*)[a]	Baseline demographic information[b]	Treatment condition (*n*)[a]	Comparison condition(s) (*n*)[a]	Mediator(s) tested	Outcome(s) tested[c]	Assessment from baseline	Method of mediation analysis
Batink *et al.* (2013)	Adults with residual depressive symptoms & ≥3 depressive episodes (59)	Mean age = 45.3, 81.4% female	MBCT + TAU (29)	TAU (30)	KIMS, PSWQ, RSS	*IDS-SR*, PSWQ, RSS, PA, *NA*	Baseline, 8 weeks	Sobel test
Batink *et al.* (2013)	Adults with residual depressive symptoms & ≤2 depressive episodes (71)	Mean age = 42.8, 70.4% female	MBCT + TAU (35)	TAU (36)	KIMS, PSWQ, RSS	*IDS-SR*, PSWQ, RSS, PA, *NA*	Baseline, 8 weeks	Sobel test
Bieling *et al.* (2012)	Adults with MDD, ≥2 depressive episodes, treated with mADM (47)	Mean age = 44, 58% female	MBCT + medication taper (15)	mADM (17), medication taper + placebo (15)	TMS, EQ	HRSD	Baseline, 8 weeks, 8 months	Kraemer et al.'s framework for RCTs
Bränström *et al.* (2010)	Adult cancer patients (71)	Mean age = 51.8, 98.6% female	MBSR without the retreat (32)	Waitlist (39)	FFMQ	*PSS*, PSOM, *IES-avoidance*	Baseline, 3 months	Causal-steps
Britton *et al.* (2012)	Adults with recurrent depression (45)	Mean age = 46.7, 85.8% female	MBCT (26)	Waitlist (19)	STAI with TSST (emotional reactivity)	BDI-II	Baseline, 8 weeks	Causal-steps, Bootstrapping
Crane *et al.* (2012)	Adults with MDD or residual symptoms & history of suicidality & ≥3 depressive episodes (27)	Mean age = 41.9 (26-64), 66.6% female	MBCT (14)	TAU (13)	AMT	Goal specificity, *perceived likelihood of attaining goals*	Baseline, 8 weeks	Causal-steps, Sobel test
Keng *et al.* (2012)	Non-clinical sample (41)	Mean age = 46.3 (21-87), 84% female	MBSR (20)	Waitlist (21)	FFMQ, SCS	ACS, PSWQ, DERS, *SAES*	Baseline, 8 weeks	Bootstrapping
Kuyken *et al.* (2010)	Adults with recurrent depression & ≥3 depressive episodes (108)	Mean age = 49.5, 77.5% female	MBCT + medication taper (51)	mADM (57)	KIMS, SCS, DAS	HRSD at 15 months	Baseline, 8 weeks, 15 months	Kraemer et al.'s framework for RCTs
Lengacher *et al.* (2014)	Adult female breast cancer survivors (82)	Mean age = 57.2 +/- 9.2, 100% female	6-week adapted MBSR (40)	Usual care (42)	CARS-30	*STAI, PSS*	Baseline, 6 weeks	Sobel test

McManus *et al.* (2012)	Adults with hypochondriasis symptoms (74)	Mean age = 42.6, 78.3% female	MBCT + US (36)	US (38)	FFMQ	SHAI	Baseline, 8 weeks, 12 months	Bootstrapping
Nyklícek & Kuijpers (2008)	Adults with symptoms of distress (57)	Mean age = 46, 66.6% female	MBSR (29)	Waitlist (28)	MAAS + KIMS	*PSS, WHOQoL-Bref*, MQ	Baseline, 8 weeks	Causal-steps with ANCOVA
Nyklícek *et al.* (2013)	Adults with symptoms of distress (146)	Mean age = 46.1 (21-66), 69% female	MBSR (72)	Waitlist (74)	MAAS + KIMS	*NA*, SI	Baseline, 8 weeks	Causal-steps with ANCOVA, Bootstrapping
Shahar *et al.* (2010)	Adults with residual depressive symptoms, partially remitted (45)	Mean age = 46.6 (24-64), 84.4% female	MBCT (26)	Waitlist (19)	MAAS, Brooding and Reflection subscales of RRS	BDI-II	Baseline, 8 weeks	Causal-steps, Bootstrapping
van Aalderen *et al.* (2012)	Adults with ≥3 depressive episodes (205)	Mean age = 47.5, 70.7% female	MBCT + TAU (102)	TAU (103)	RSS, PSWQ, KIMS	HAMD	Baseline, 8 weeks	Causal-steps, Bootstrapping
Vøllestad *et al.* (2011)	Adults with heterogeneous anxiety disorders (76)	Mean age = 42.5, 67.1% female	MBSR (39)	Waitlist (37)	FFMQ	*BAI*, PSWQ, *BDI-II, STAI-T*	Baseline, 8 weeks	Causal-steps, Bootstrapping

Note. ACS = Affective Control Scale (Williams, Chambless, & Ahrens, 1997); AMT = Autobiographical Memory Test (Williams & Broadbent, 1986); BAI = Beck Anxiety Inventory (Beck & Steer, 1993); BDI-II = Beck Depression Inventory-II (Beck, Steer, & Brown, 1996); CARS-30 = 30-item Concerns About Recurrence Scale (Vickberg, 2003); DAS = Dysfunctional Attitude Scale, Versions A and B (Weissman & Beck, 1978); DERS = Difficulties in Emotion Regulation Scale (Gratz & Roemer, 2004); EQ = Experiences Questionnaire (Fresco et al., 2007); FFMQ = Five-Facet Mindfulness Questionnaire (Baer et al., 2006); HAMD = Hamilton Rating Scale for Depression (Hamilton, 1960); HSRD = Hamilton Rating Scale for Depression (Williams, 1988); IDS-SR = Inventory of Depressive Symptoms-Self-Rating (Rush, Gullion, Bascro, & Jarrett, 1996); IES-avoidance = avoidance subscale from The Impact of Event Scale Revised (posttraumatic stress) (Horowitz, Wilner, & Alvarez, 1979); KIMS = Kentucky Inventory of Mindfulness Skills (Baer, Smith, & Allen, 2004); MAAS = Mindful Attention and Awareness Scale (Brown & Ryan, 2003); mADM = maintenance antidepressant medication; MANCOVA = multivariate analysis of covariance; MBCT = mindfulness-based cognitive therapy; MBSR = mindfulness-based stress reduction; MDD = major depressive disorder; MQ = Maastricht Vital Exhaustion Questionnaire (Appels, & Mulder, 1989); NA = negative affect; PA = positive affect; PSOM = Positive States of Mind Questionnaire (Adler, Horowitz, Garcia, & Moyer, 1998); PSS = Perceived Stress Scale (Cohen & Williamson, 1988); PSWQ = Penn State Worry Questionnaire (Meyer, Miller, Metzger, & Borkovec, 1990); RCT = randomised controlled trial; RRS = Ruminative Response Scale (Treynor, Gonzalez, & Nolen-Hoeksema, 2003); RRQ = Rumination and Reflection Questionnaire (Trapnell & Campbell, 1999); RSS = Rumination on Sadness Scale (Raes, Hermans, & Eelen, 2003); SAES = Spielberger Anger Expression Scale (Spielberger et al., 1985); SCS = Self-Compassion Scale (Neff, 2003b); SHAI = Short Health Anxiety Inventory (Salkovskis, Rimes, Warwick, & Clark, 2002); SI = social inhibition; STAI = State-Trait Anxiety Inventory (Speilberger, Gorsuch, & Lushene, 1970); STAI-T = State-Trait Anxiety Inventory-Trait; TAU = treatment as usual; TMS = Toronto Mindfulness Scale (Lau et al., 2006); TSST = Trier Social Stress Test (Kirschbaum, Pirke, & Hellhammer, 1993); US = unrestricted services; WHOQoL-Bref = World Health Organisation Quality of Life Questionnaire – Short Form (WHOQOL Group, 1998).

[a] Sample size refers to participants included in mediation analyses.

[b] Demographic information (mean age, % female) refers to characteristics of the whole sample at baseline.

[c] For studies which tested more than one outcome in their mediation analyses, those which fit the inclusion criteria and are reported on are italicised.

Table 4.
Descriptions of included quasi-experimental studies

Study	Sample (*N*)[a]	Demographic information [b]	Treatment condition (*n*)[a]	Comparison condition(s) (*n*)[a]	Mediator(s) tested	Outcome(s) tested [c]	Assessment from baseline	Method of mediation analysis
Bergen-Cico & Cheon (2013)	Non-clinical sample (202)	Mean age = 23.17 (19-61), 73% female	MBSR (108)	Psychoeducation and interactive response exercises (94)	KIMS, SCS	*STAI-T*, SCS	Baseline, 6 weeks, 11 weeks	Path analysis, Causal-steps, Sobel test
Heeren & Philippot (2011)	Non-clinical sample (41)	Mean age = 47.9 (22-75), 65.9% female	MBCT with greater focus on psychoeducation (29)	Waitlist (12)	Mini-CERTS	*GSI of SCL-90-R*	Baseline, 8 weeks	Causal-steps, Sobel test
Labelle (2012)	Adult cancer patients (136)	Mean age = 52.7, 80.1% female	MBSR (75)	Waitlist (61)	MAAS, FFMQ, RRQ, PSWQ, AAQ	*CSOSI, POMS*, RRQ, PSWQ, AAQ	Baseline, 8 weeks	Causal-steps, Kraemer et al.'s framework for RCTs, Bootstrapping
Labelle *et al.* (2010)	Adult female cancer patients who have completed all treatments (77)	Mean age = 53.1, 100% female	MBSR slightly adapted for cancer recovery (46)	Waitlist (31)	MAAS, RRQ	CESD-10	Baseline, 8 weeks	Causal-steps, Bootstrapping
Raes *et al.* (2009)	Unselected adult sample (39)	Mean age = 41.9, 84.1% female	MBCT (18)	Waitlist (21)	KIMS	LEIDS-R	Baseline, 8 weeks	Bootstrapping

Note. AAQ = Acceptance and Action Questionnaire (Hayes et al., 2004); CESD-10 = Center for Epidemiological Studies Depression Inventory-10 (Andresen, Malmgren, Carter, & Patrick, 1994); CSOSI = Calgary Symptoms of Stress Inventory (Leckie & Thompson, 1979); FFMQ = Five-Facet Mindfulness Questionnaire (Baer et al., 2006); GSI of SCL-90-R = Global Score Index of the Psychopathological Symptom Checklist (Derogatis, 1977); KIMS = Kentucky Inventory of Mindfulness Skills (Baer, Smith, & Allen, 2004); LEIDS-R = Leiden Index of Depression Sensitivity-Revised (Van der Does, 2002); MAAS = Mindful Attention and Awareness Scale (Brown & Ryan, 2003); MBCT = mindfulness-based cognitive therapy; MBSR = mindfulness-based stress reduction; Mini-CERTS = Cambridge-Exeter Rumination Thinking Scale Short-Form (Barnard, Watkins, Mackintosh, & Nimmo-Smith, 2007); POMS = Profile of Mood States (McNair, Lorr, & Droppelman, 1971); PSWQ = Penn State Worry Questionnaire (Meyer, Miller, Metzger, & Borkovec, 1990); RRQ = Rumination and Reflection Questionnaire (Trapnell & Campbell, 1999); SCS = Self-Compassion Scale (Neff, 2003b); SEM = structural equation modeling; STAI-T = State-Trait Anxiety Inventory-Trait (Speilberger, Gorsuch, & Lushene, 1970).

[a] Sample size refers to participants included in mediation analyses.

[b] Demographic information (mean age, % female) refers to characteristics of the whole sample at baseline.

[c] For studies which tested more than one outcome in their mediation analyses, those which fit the inclusion criteria and are reported on are italicised.

Appendix C: Certificate of Ethical Approval for Chapter 3

University of Sussex

Certificate of Approval	
Reference Number:	ER/JG252/3
Title Of Project:	How does an online mindfulness-based self-help intervention improve mental health?
Principal Investigator (PI):	Clara Strauss
Student:	Jenny Gu
Collaborators:	
Duration Of Approval:	5 months
Expected Start Date:	19-Jan-2015
Date Of Approval:	08-Feb-2015
Approval Expiry Date:	30-Jun-2015
Approved By:	Richard de Visser
Name of Authorised Signatory:	Richard de Visser
Date:	08-Feb-2015

*NB. If the actual project start date is delayed beyond 12 months of the expected start date, this Certificate of Approval will lapse and the project will need to be reviewed again to take account of changed circumstances such as legislation, sponsor requirements and University procedures.

Please note and follow the requirements for approved submissions:

Amendments to protocol

* Any changes or amendments to approved protocols must be submitted to the C-REC for authorisation prior to implementation.

Feedback regarding the status and conduct of approved projects

* Any incidents with ethical implications that occur during the implementation of the project must be reported immediately to the Chair of the C-REC.

Feedback regarding any adverse and unexpected events

* Any adverse (undesirable and unintended) and unexpected events that occur during the implementation of the project must be reported to the Chair of the Social Sciences C-REC. In the event of a serious adverse event, research must be stopped immediately and the Chair alerted within 24 hours of the occurrence.

For Life Sciences and Psychology projects

* The principal investigator is required to provide a brief annual written statement to the committee, indicating the status and conduct of the approved project. These reports will be reviewed at the annual meeting of the committee. A statement by the PI to the C-REC indicating the status and conduct of the approved project will be required on the Approval Expiration Date as stated above.

Appendix D: Supplementary Materials for Chapter 4

Table 1.
Unstandardised and standardised loadings for the FFMQ-39 four-factor hierarchical[a] *pre-MBCT model*

Path	Unstandardised estimate (*SE*)	Standardised estimate
Factor loadings		
Describing <- Mindfulness	1.31 (.24)***	.63
Acting with awareness <- Mindfulness	1.38 (.25)***	.72
Non-judging <- Mindfulness	1.05 (.21)***	.53
Non-reactivity <- Mindfulness	1.00	.63
Item loadings		
Parcel 1 <- Acting with awareness	1.00	.89
Parcel 2 <- Acting with awareness	0.91 (.06)***	.84
Parcel 3 <- Acting with awareness	0.83 (.06)***	.75
Parcel 1 <- Non-judging	1.00	.85
Parcel 2 <- Non-judging	0.92 (.07)***	.80
Parcel 3 <- Non-judging	1.14 (.09)***	.80
Parcel 1 <- Describing	1.00	.87
Parcel 2 <- Describing	0.98 (.06)***	.87
Parcel 3 <- Describing	1.19 (.06)***	.90
Parcel 1 <- Non-reactivity	1.00	.77
Parcel 2 <- Non-reactivity	1.20 (.11)***	.78
Parcel 3 <- Non-reactivity	1.33 (.11)***	.84

Note. FFMQ = Five-Facet Mindfulness Questionnaire; MBCT = Mindfulness-Based Cognitive Therapy; SE = standard error.
[a] Four-factor hierarchical refers to the model in which the facets describe, acting with awareness, non-judging, and non-reactivity (without the observing facet) loaded onto an overall mindfulness factor.
***$p < .001$.

Table 2.
Unstandardised and standardised loadings for the FFMQ-39 five-factor hierarchical [a] *post-MBCT model*

Path	Unstandardised estimate (*SE*)	Standardised estimate
Factor loadings		
Observing <- Mindfulness	0.77 (.11)***	.63
Describing <- Mindfulness	0.94 (.13)***	.59
Acting with awareness <- Mindfulness	0.85 (.11)***	.65
Non-judging <- Mindfulness	0.92 (.12)***	.66
Non-reactivity <- Mindfulness	1.00	.86
Item loadings		
Parcel 1 <- Observing	1.00	.82
Parcel 2 <- Observing	1.06 (.08)***	.92
Parcel 3 <- Observing	0.99 (.09)***	.67
Parcel 1 <- Describing	1.00	.91
Parcel 2 <- Describing	0.92 (.05)***	.89
Parcel 3 <- Describing	1.07 (.05)***	.90
Parcel 1 <- Non-reactivity	1.00	.88
Parcel 2 <- Non-reactivity	1.03 (.08)***	.75
Parcel 3 <- Non-reactivity	1.17 (.08)***	.82
Parcel 1 <- Non-judging	1.00	.85
Parcel 2 <- Non-judging	1.08 (.07)***	.84
Parcel 3 <- Non-judging	1.12 (.08)***	.83
Parcel 1 <- Acting with awareness	1.00	.89
Parcel 2 <- Acting with awareness	0.92 (.06)***	.83
Parcel 3 <- Acting with awareness	0.95 (.06)***	.84

Note. FFMQ = Five-Facet Mindfulness Questionnaire; MBCT = Mindfulness-Based Cognitive Therapy; SE = standard error.
[a] Five-factor hierarchical refers to the model in which all five facets loaded onto an overall mindfulness factor.
***$p < .001$.

Table 3.
Unstandardised and standardised loadings for the FFMQ-15 four-factor hierarchical[a] *pre-MBCT model*

Path	Unstandardised estimate (*SE*)	Standardised estimate
Factor loadings		
Describing <- Mindfulness	0.93 (.28)***	.41
Acting with awareness <- Mindfulness	0.56 (.17)***	.59
Non-judging <- Mindfulness	1.64 (.42)***	.73
Non-reactivity <- Mindfulness	1.00	.58
Item loadings		
Item #8 <- Acting with awareness	1.00	.41
Item #34 <- Acting with awareness	2.02 (.36)***	.71
Item #38 <- Acting with awareness	2.33 (.43)***	.86
Item #14 <- Non-judging	1.00	.74
Item #10 <- Non-judging	0.74 (.09)***	.62
Item #30 <- Non-judging	1.03 (.11)***	.79
Item #2 <- Describing	1.00	.80
Item #16 <- Describing	0.76 (.08)***	.64
Item #27 <- Describing	1.11 (.11)***	.84
Item #19 <- Non-reactivity	1.00	.63
Item #29 <- Non-reactivity	1.29 (.20)***	.85
Item #33 <- Non-reactivity	0.69 (.12)***	.46

Note. FFMQ = Five-Facet Mindfulness Questionnaire; MBCT = Mindfulness-Based Cognitive Therapy; SE = standard error.
[a] Four-factor hierarchical refers to the model in which the facets describe, acting with awareness, non-judging, and non-reactivity (without the observing facet) loaded onto an overall mindfulness factor.
***$p < .001$.

Table 4.
Unstandardised and standardised loadings for the FFMQ-15 five-factor hierarchical [a] *post-MBCT model*

Path	Unstandardised estimate (*SE*)	Standardised estimate
Factor loadings		
Observing <- Mindfulness	0.72 (.16)***	.55
Describing <- Mindfulness	0.86 (.17)***	.51
Acting with awareness <- Mindfulness	0.59 (.13)***	.65
Non-judging <- Mindfulness	1.06 (.19)***	.72
Non-reactivity <- Mindfulness	1.00	.73
Item loadings		
Item #6 <- Observing	1.00	.64
Item #11 <- Observing	1.00 (.15)***	.56
Item #15 <- Observing	1.11 (.16)***	.78
Item #2 <- Describing	1.00	.78
Item #16 <- Describing	0.95 (.09)***	.76
Item #27 <- Describing	1.08 (.10)***	.81
Item #19 <- Non-reactivity	1.00	.71
Item #29 <- Non-reactivity	1.04 (.11)***	.82
Item #33 <- Non-reactivity	0.87 (.10)***	.68
Item #14 <- Non-judging	1.00	.72
Item #10 <- Non-judging	0.95 (.11)***	.63
Item #30 <- Non-judging	1.24 (.12)***	.89
Item #8 <- Acting with awareness	1.00	.49
Item #34 <- Acting with awareness	1.48 (.23)***	.70
Item #38 <- Acting with awareness	1.78 (.28)***	.84

Note. FFMQ = Five-Facet Mindfulness Questionnaire; MBCT = Mindfulness-Based Cognitive Therapy; SE = standard error.
[a] Five-factor hierarchical refers to the model in which all five facets loaded onto an overall mindfulness factor.
***$p < .001$.

Appendix E: Supplementary Materials for Chapter 6

Pool of 80 Compassion Items from Stage 1

The origin of each item is given in brackets: CLS = Compassionate Love Scale (Sprecher & Fehr, 2005); CS-M = Martins et al.'s (2013) Compassion Scale; CS-P = Pommier's (2010) Compassion Scale; EC = generated through expert consultation; RCS = Relational Compassion Scale (Hacker, 2008).
[a] Only item stems (50% or less of the full items) are given for items from CS-P. For full items, please refer to Pommier (2010).

1. I notice when people are upset…[a] (CS-P).
2. If I see someone going through a difficult…[a] (CS-P).
3. I try to put myself in other people's shoes when they are in trouble (CLS).
4. My heart goes out…[a] (CS-P).
5. When other people are upset, I try to be warm, sensitive and sympathetic to them (RCS).
6. I feel uncomfortable being around people who are suffering (EC).
7. If a person needs help, I would do almost anything I could to help him or her (CLS).
8. I know that everyone feels down…[a] (CS-P).
9. I feel happy when I see that other people are happy (CLS).
10. If someone does something disagreeable I find it hard to be kind to them when they are in need (EC).
11. It is easy for me to feel the pain (and joy) experienced by other people (CLS).
12. It is important to me to recognize that…[a] (CS-P).
13. When other people are emotionally upset I treat them with kindness and care (RCS).
14. I pay careful attention when…[a] (CS-P).
15. When I hear about a person going through a difficult time, I feel a great deal of compassion for him or her (CLS).
16. I find it difficult to notice when people are upset (EC).
17. I am caring towards others when they are distressed (RCS).
18. I find it difficult to understand why other people get upset (EC).
19. When someone is suffering it can be hard to help them because it is so upsetting (EC).
20. Others can trust that I will be there for them if they need me (CLS).
21. When I'm busy, I do not notice other people's emotions (EC).
22. If I am upset by another's suffering, I am able to tolerate my own emotions and not get carried away by them (EC).
23. If given the opportunity, I am willing to make sacrifices in order to let other people achieve their goals in life (CLS).
24. I think that sometimes people get upset over nothing (EC).
25. When someone is upset I help them without judging them (EC).
26. Despite my differences with others, I know…[a] (CS-P).
27. When someone is troubled, I feel extreme tenderness and caring (CLS).
28. If someone is in distress or trouble, I wait for other people to respond first (EC).
29. I can't really connect with…[a] (CS-P).
30. I want to spend time with others so that I can find ways to help enrich their lives (CLS).
31. I believe that suffering is just…[a] (CS-P).
32. When people tell me about their problems, I…[a] (CS-P).
33. I don't concern myself…[a] (CS-P).
34. I show understanding and caring towards others (RCS).
35. I very much wish to be kind and good to other people (CLS).
36. I find it hard to understand other people's problems (RCS).
37. I am moved by other people's suffering (EC).
38. I would be willing to give money to help someone in need of financial help (CS-M).
39. Thinking of other people's suffering as part of common human experience doesn't come easily to me (EC).
40. I like to be there…[a] (CS-P).
41. I find it difficult to recognise how other people are feeling if they don't say anything (EC).
42. I am interested to understand others' experiences and emotions (RCS).
43. I would not be willing to help someone who I find unpleasant (EC).
44. I am attuned to other peoples' feelings (RCS).
45. When other people are emotionally upset I try to see their thoughts and feelings as valid (RCS).

46. I would be willing to share my home with someone in need who poses no threat (CS-M).
47. I feel detached from others…[a] (CS-P).
48. When I see someone suffering I think about how I too could be in their position (EC).
49. I can understand why other people get upset (EC).
50. One of the activities that provides me with the most meaning to my life is helping others (CLS).
51. When people cry in front of…[a] (CS-P).
52. I accept other people even when they do things I think are wrong (CLS).
53. I listen patiently when…[a] (CS-P).
54. Even if it is upsetting hearing about someone's problems, I do my best to listen (EC).
55. When I see people feeling sad, I feel a need to reach out to them (CLS).
56. I don't feel emotionally…[a] (CS-P).
57. It is hard for me to relate to others when I see them suffering (EC).
58. When others are feeling troubled, I…[a] (CS-P).
59. I feel compassion for other people (CLS).
60. When I see someone feeling upset I feel so overwhelmed by my emotions that I find it difficult to help them (EC).
61. I am accepting of people who are in need whatever their experiences (EC).
62. When others feel sadness…[a] (CS-P).
63. I find it easy to recognise when someone is suffering or in need (EC).
64. When people talk about their problems…[a] (CS-P).
65. I try to avoid people who…[a] (CS-P).
66. I am empathetic towards others when they make a mistake (RCS).
67. I don't know what to do when other people are distressed (RCS).
68. I can understand how people are feeling even if I do not identify with their experiences (EC).
69. I am sensitive to the wellbeing of others (RCS).
70. I get carried away by my own emotional response to other people's problems or suffering (EC).
71. I notice when someone is different from how they usually are (EC).
72. I would be willing to give my time freely to work for someone who needs my skills but cannot afford to pay me (CS-M).
73. I try to understand rather than judge people (CLS).
74. I am cold to…[a] (CS-P).
75. I tune out when people…[a] (CS-P).
76. If someone is suffering I go out of my way to help them if I can (EC).
77. I have tender feelings towards others when they seem to be in need (CLS).
78. When I see someone feeling down, I…[a] (CS-P).
79. I spend a lot of time concerned about the well-being of other people (CLS).
80. If someone is suffering I feel a natural response to want to help (EC).

Appendix E: Certificate of Ethical Approval for Chapter 6

University of Sussex

Certificate of Approval	
Reference Number:	ER/JG252/2
Title Of Project:	Developing a measure of compassion towards others and examining the roles of self-compassion and other-compassion as mechanisms of mindfulness
Principal Investigator (PI):	Kate Cavanagh
Student:	Jenny Gu
Collaborators:	
Duration Of Approval:	5 months
Expected Start Date:	19-Jan-2015
Date Of Approval:	07-Feb-2015
Approval Expiry Date:	30-Jun-2015
Approved By:	Richard de Visser
Name of Authorised Signatory:	Richard de Visser
Date:	07-Feb-2015

*NB. If the actual project start date is delayed beyond 12 months of the expected start date, this Certificate of Approval will lapse and the project will need to be reviewed again to take account of changed circumstances such as legislation, sponsor requirements and University procedures.

Please note and follow the requirements for approved submissions:

Amendments to protocol

* Any changes or amendments to approved protocols must be submitted to the C-REC for authorisation prior to implementation.

Feedback regarding the status and conduct of approved projects

* Any incidents with ethical implications that occur during the implementation of the project must be reported immediately to the Chair of the C-REC.

Feedback regarding any adverse and unexpected events

* Any adverse (undesirable and unintended) and unexpected events that occur during the implementation of the project must be reported to the Chair of the Social Sciences C-REC. In the event of a serious adverse event, research must be stopped immediately and the Chair alerted within 24 hours of the occurrence.

For Life Sciences and Psychology projects

* The principal investigator is required to provide a brief annual written statement to the committee, indicating the status and conduct of the approved project. These reports will be reviewed at the annual meeting of the committee. A statement by the PI to the C-REC indicating the status and conduct of the approved project will be required on the Approval Expiration Date as stated above.

Appendix F: Supplementary Materials for Chapter 7

Table 1.
Standardised loadings of compassion for others items on each factor in a sample of 932 healthcare staff (Stage 2)

	Standardised loading	*SE*
Recognising suffering		
*I notice when others are feeling distressed.	0.826	0.02
*I recognise when other people are feeling distressed without them having to tell me.	0.799	0.02
*I recognise signs of suffering in others.	0.797	0.02
*I'm quick to notice early signs of distress in others.	0.780	0.02
Even if I'm busy, I can still tell when other people are upset.	0.770	0.02
I can tell when someone else is holding back tears.	0.696	0.02
I recognise when other people first start feeling upset.	0.689	0.03
I can tell when someone else is pretending to be fine when they're not.	0.672	0.02
I'm better than most people at recognising when others are struggling emotionally.	0.647	0.03
It takes me a while to notice if someone is upset about something. (negatively phrased)	0.498	0.04
I misjudge how others are feeling. (negatively phrased)	0.489	0.03
I'm unaware of other people's distress. (negatively phrased)	0.173	0.04
Understanding the universality of suffering		
*I understand that everyone experiences suffering at some point in their lives.	0.872	0.01
*I understand that feeling upset at times is part of human nature.	0.851	0.02
*Like me, I know that other people also experience struggles in life.	0.811	0.02
*I know that we can all feel upset at times when we are wronged.	0.794	0.02
I believe that feeling low at times is part of being human.	0.772	0.02
I know that we can all feel distressed when things don't go well in our lives.	0.768	0.02
I understand that experiences of suffering are shared by everybody.	0.754	0.02
I know that everyone has ups and downs in their lives.	0.673	0.03
I believe that no matter how different people are, we all experience difficulties in life.	0.672	0.03
I know that I'm not alone in experiencing suffering.	0.641	0.03
I believe that suffering is a shared human experience.	0.608	0.02
When someone else is going through a hard time, I can understand and relate to their experience.	0.355	0.03
Feeling for the person suffering		
*When someone is going through a difficult time, I feel kindly towards them.	0.769	0.02
*When I hear about bad things happening to other people, I feel concern for their wellbeing.	0.707	0.03
*When someone is upset, I try to tune in to how they're feeling.	0.688	0.02
*I'm sensitive to other people's distress.	0.682	0.03
Even if I don't like someone, I can feel warmly towards them when they're in distress.	0.638	0.03
I'm emotionally distant from others when they're upset. (negatively phrased)	0.557	0.03
I feel indifferent towards other people's struggles. (negatively phrased)	0.512	0.04
I feel emotionally detached from other people's problems. (negatively phrased)	0.472	0.04
Tolerating uncomfortable feelings		
*When someone else is upset, I try to stay open to their feelings rather than avoid them.	0.750	0.03
*I stay with and listen to other people when they're upset even if it's hard to bear.	0.707	0.03
*I connect with the suffering of others without judging them.	0.621	0.04
*When someone else is upset, I can be there for them without feeling overwhelmed by their distress.	0.437	0.06
I can stay with uncomfortable feelings in response to other people's distress.	0.383	0.04

I connect with other people's distress without letting it overwhelm me.	0.378	0.06
I can be around someone who is upset without letting their distress take over me.	0.367	0.06
When others are upset, I feel so uncomfortable that I'm unable to be around them. (negatively phrased)	0.343	0.04
I feel overwhelmed by other people's problems. (negatively phrased)	0.204	0.05
If I feel uncomfortable in response to someone else's distress, I don't let myself get carried away by my feelings.	0.204	0.05
Acting or being motivated to act to alleviate suffering		
*When others are struggling, I try to do things that would be helpful.	0.823	0.02
*When someone is going through a difficult time, I try to look after them.	0.815	0.02
*When I see someone in need, I try to do what's best for them.	0.806	0.02
*When I see that someone is upset, I do my best to take care of them.	0.805	0.02
If someone is suffering, I try to do something helpful to reduce their distress.	0.800	0.02
I try to help people feel better when they are distressed, even if I can't do anything about the problem.	0.790	0.02
When someone is going through a hard time, I try to do what's best for them.	0.726	0.02
I try to prioritise the needs of others when they're facing difficulties.	0.690	0.02
When others are feeling down, I think of ways I can make things better for them.	0.685	0.02
I think of useful ways I can help people when they're struggling.	0.664	0.03
If someone looks troubled, I stop and ask if there's anything I can do.	0.622	0.02
I'm drawn to helping those in need.	0.597	0.03
I'm dismissive of other people's difficulties and problems. (negatively phrased)	0.487	0.04
I go easy on other people when they're feeling upset.	0.466	0.04
When others are going through a difficult time, I leave them to deal with their own problems. (negatively phrased)	0.431	0.04

Note. Items preceded by an asterisk indicate ones which were selected for inclusion in the SOCS-O. All standardised loadings were significant ($p < .001$). Standardised loadings are reported to three decimal places to facilitate selection of the highest loading items.

Table 2.
Standardised loadings of compassion for the self items on each factor in a sample of 947 healthcare staff (Stage 2)

	Standardised loading	*SE*
Recognising suffering		
*I'm quick to notice early signs of distress in myself.	0.833	0.01
*I notice when I'm feeling distressed.	0.833	0.01
*I'm good at recognising when I'm feeling distressed.	0.778	0.02
*I recognise signs of suffering in myself.	0.773	0.02
Even if I'm busy, I can still tell when I'm upset.	0.713	0.02
Compared to most people, I'm better at recognising when I'm struggling emotionally.	0.657	0.02
I recognise when I first start feeling upset.	0.656	0.02
It takes me a while to notice if I'm upset about something. (negatively phrased)	0.510	0.03
I can tell when I'm holding back tears.	0.504	0.03
I can tell when I'm pretending to be fine when I'm not.	0.488	0.03
I'm unaware of my own distress. (negatively phrased)	0.377	0.04
Understanding the universality of suffering		
*I understand that everyone experiences suffering at some point in their lives.	0.873	0.02
*Like me, I know that other people also experience struggles in life.	0.858	0.01
*I understand that feeling upset at times is part of human nature.	0.842	0.02
*I know that we can all feel distressed when things don't go well in our lives.	0.802	0.02
I understand that experiences of suffering are shared by everybody.	0.802	0.03
I know that we can all feel upset at times when we are wronged.	0.790	0.02
I know that everyone has ups and downs in their lives.	0.735	0.02
I believe that feeling low at times is part of being human.	0.734	0.02
I believe that no matter how different people are, we all experience difficulties in life.	0.678	0.03
I believe that suffering is a shared human experience.	0.641	0.02
I know that I'm not alone in experiencing suffering.	0.615	0.03
When I'm going through a hard time, I take comfort in knowing that we all go through similar experiences.	0.388	0.03
Feeling for the person suffering		
*When I'm going through a difficult time, I feel kindly towards myself.	0.859	0.01
*When bad things happen to me, I feel caring towards myself.	0.818	0.02
*Even when I'm disappointed with myself, I can feel warmly towards myself when I'm in distress.	0.783	0.02
*When I'm upset, I try to tune in to how I'm feeling.	0.652	0.03
I'm sensitive to my own distress.	0.510	0.04
I feel indifferent towards my own struggles. (negatively phrased)	0.402	0.04
I block out my feelings when I'm upset. (negatively phrased)	0.375	0.04
I feel emotionally detached from my own problems. (negatively phrased)	0.204	0.04
Tolerating uncomfortable feelings		
*When I'm upset, I try to stay open to my feelings rather than avoid them.	0.718	0.03
*I connect with my own suffering without judging myself.	0.716	0.02
*When I'm upset, I can let the emotions be there without feeling overwhelmed.	0.687	0.03
*I connect with my own distress without letting it overwhelm me.	0.684	0.03
I face my own feelings when I'm upset even if it's hard to bear.	0.668	0.03
I can be upset without letting my feelings take over me.	0.651	0.03
I'm able to tolerate uncomfortable feelings when I'm distressed.	0.608	0.03
I feel overwhelmed by my own problems. (negatively phrased)	0.601	0.03
When I'm in distress, I allow myself to experience the emotions as they are, even if they're uncomfortable.	0.589	0.03
When I'm upset, I feel so uncomfortable that I push the feelings away. (negatively phrased)	0.539	0.03
I'm judgemental of myself when I'm going through a hard time. (negatively phrased)	0.487	0.04

I overreact to my own problems. (negatively phrased)	0.486	0.03
Acting or being motivated to act to alleviate suffering		
*When I'm upset, I try to do what's best for myself.	0.867	0.01
*When I'm going through a difficult time, I try to look after myself.	0.866	0.01
*I try to make myself feel better when I'm distressed, even if I can't do anything about the cause.	0.848	0.01
*When I'm upset, I do my best to take care of myself.	0.846	0.01
When I'm struggling, I try to do things that would be helpful.	0.837	0.01
When I'm going through a hard time, I try to do what's best for myself.	0.799	0.02
If I'm suffering, I try to do something helpful to reduce my distress.	0.787	0.02
I feel a desire to help myself when I'm in need.	0.783	0.02
When I feel troubled, I stop to ask myself what I could do that might help.	0.732	0.02
I go easy on myself when I'm feeling upset.	0.712	0.02
I think of useful ways I can help myself when I'm struggling.	0.707	0.02
I try to prioritise my own needs when I'm facing difficulties.	0.692	0.02
When I'm feeling down, I think of ways I can make things better for myself.	0.634	0.03
When I'm going through a difficult time, I can be hard on myself. (negatively phrased)	0.516	0.03
I'm dismissive of my own difficulties and problems. (negatively phrased)	0.490	0.03

Note. Items preceded by an asterisk indicate ones which were selected for inclusion in the SOCS-S. All standardised loadings were significant ($p < .001$). Standardised loadings are reported to three decimal places to facilitate selection of the highest loading items.

Table 3.
Standardised item loadings for the compassion for others five-factor hierarchical model in both validation samples (Stages 3 and 4)

	1,242 healthcare staff (Stage 3)	371 students (Stage 4)
Recognising suffering		
I notice when others are feeling distressed.	0.88 (0.01)**	0.83 (0.02)**
I recognise when other people are feeling distressed without them having to tell me.	0.78 (0.02)**	0.76 (0.04)**
I recognise signs of suffering in others.	0.82 (0.01)**	0.70 (0.04)**
I'm quick to notice early signs of distress in others.	0.83 (0.01)**	0.82 (0.02)**
Understanding the universality of suffering		
I understand that everyone experiences suffering at some point in their lives.	0.90 (0.01)**	0.84 (0.03)**
I understand that feeling upset at times is part of human nature.	0.89 (0.01)**	0.86 (0.03)**
Like me, I know that other people also experience struggles in life.	0.85 (0.02)**	0.81 (0.04)**
I know that we can all feel upset at times when we are wronged.	0.81 (0.02)**	0.77 (0.04)**
Feeling for the person suffering		
When someone is going through a difficult time, I feel kindly towards them.	0.72 (0.14)**	0.66 (0.06)**
When I hear about bad things happening to other people, I feel concern for their wellbeing.	0.70 (0.02)**	0.58 (0.05)**
When someone is upset, I try to tune in to how they're feeling.	0.69 (0.02)**	0.69 (0.03)**
I'm sensitive to other people's distress.	0.73 (0.02)**	0.60 (0.05)**
Tolerating uncomfortable feelings		
When someone else is upset, I try to stay open to their feelings rather than avoid them.	0.77 (0.02)**	0.75 (0.04)**
I stay with and listen to other people when they're upset even if it's hard to bear.	0.72 (0.02)**	0.70 (0.04)**
I connect with the suffering of others without judging them.	0.71 (0.02)**	0.56 (0.06)**
When someone else is upset, I can be there for them without feeling overwhelmed by their distress.	0.42 (0.03)**	0.17 (0.06)*
Acting or being motivated to act to alleviate suffering		
When others are struggling, I try to do things that would be helpful.	0.84 (0.02)**	0.76 (0.03)**
When someone is going through a difficult time, I try to look after them.	0.87 (0.01)**	0.82 (0.03)**
When I see someone in need, I try to do what's best for them.	0.83 (0.01)**	0.76 (0.03)**
When I see that someone is upset, I do my best to take care of them.	0.82 (0.01)**	0.75 (0.03)**

Standard errors are given in parentheses. ** $p < .001$, * $p < .01$.

Table 4.
Standardised loadings of factors to an overall compassion factor in the five-factor hierarchical model in both validation samples (Stages 3 and 4)

	Compassion for others		Compassion for the self	
	1,242 healthcare staff (Stage 3)	371 students (Stage 4)	1,216 healthcare staff (Stage 3)	371 students (Stage 4)
Recognising suffering	0.78 (0.02)*	0.62 (0.05)*	0.59 (0.03)*	0.33 (0.06)*
Understanding the universality of suffering	0.58 (0.03)*	0.42 (0.06)*	0.34 (0.03)*	0.26 (0.05)*
Feeling for the person suffering	0.98 (0.01)*	0.92 (0.04)*	0.99 (0.01)*	0.98 (0.02)*
Tolerating uncomfortable feelings	0.93 (0.02)*	0.91 (0.03)*	0.94 (0.01)*	0.87 (0.03)*
Acting or being motivated to act to alleviate suffering	0.92 (0.01)*	0.95 (0.03)*	0.94 (0.01)*	0.94 (0.02)*

Standard errors are given in parentheses. * $p < .001$.

Table 5.
Correlation coefficients between total scale and subscale scores on the SOCS-O and SOCS-S using available data from 1,319 healthcare staff (Stage 3)

	SOCS-O	1	2	3	4	5	SOCS-S	6	7	8	9	10
Sussex-Oxford Compassion for Others Scale (SOCS-O)	-											
1. Recognising suffering	.80*	-										
2. Understanding the universality of suffering	.70*	.41*	-									
3. Feeling for the person suffering	.88*	.66*	.49*	-								
4. Tolerating uncomfortable feelings	.84*	.61*	.48*	.68*	-							
5. Acting or being motivated to act to alleviate suffering	.87*	.64*	.47*	.77*	.71*	-						
Sussex-Oxford Compassion for the Self Scale (SOCS-S)	.40*	.33*	.39*	.32*	.37*	.25*	-					
6. Recognising suffering	.37*	.36*	.33*	.30*	.31*	.25*	.74*	-				
7. Understanding the universality of suffering	.51*	.27*	.78*	.36*	.36*	.32*	.55*	.41*	-			
8. Feeling for the person suffering	.25*	.23*	.20*	.23*	.25*	.15*	.89*	.52*	.30*	-		
9. Tolerating uncomfortable feelings	.25*	.22*	.17*	.19*	.32*	.15*	.83*	.46*	.29*	.75*	-	
10. Acting or being motivated to act to alleviate suffering	.23*	.22*	.16*	.19*	.24*	.15*	.88*	.52*	.28*	.84*	.72*	-

*$p < .001$.

Table 6.
Standardised item loadings for the compassion for the self five-factor hierarchical model in both validation samples (Stages 3 and 4)

	1,216 healthcare staff (Stage 3)	371 students (Stage 4)
Recognising suffering		
I'm quick to notice early signs of distress in myself.	0.85 (0.01)*	0.80 (0.03)*
I notice when I'm feeling distressed.	0.81 (0.02)*	0.81 (0.03)*
I'm good at recognising when I'm feeling distressed.	0.80 (0.02)*	0.77 (0.03)*
I recognise signs of suffering in myself.	0.76 (0.02)*	0.69 (0.04)*
Understanding the universality of suffering		
I understand that everyone experiences suffering at some point in their lives.	0.88 (0.02)*	0.91 (0.02)*
Like me, I know that other people also experience struggles in life.	0.86 (0.02)*	0.87 (0.03)*
I understand that feeling upset at times is part of human nature.	0.87 (0.02)*	0.85 (0.03)*
I know that we can all feel distressed when things don't go well in our lives.	0.84 (0.02)*	0.77 (0.04)*
Feeling for the person suffering		
When I'm going through a difficult time, I feel kindly towards myself.	0.85 (0.01)*	0.87 (0.02)*
When bad things happen to me, I feel caring towards myself.	0.79 (0.02)*	0.87 (0.02)*
Even when I'm disappointed with myself, I can feel warmly towards myself when I'm in distress.	0.77 (0.02)*	0.76 (0.03)*
When I'm upset, I try to tune in to how I'm feeling.	0.65 (0.02)*	0.54 (0.05)*
Tolerating uncomfortable feelings		
When I'm upset, I try to stay open to my feelings rather than avoid them.	0.69 (0.02)*	0.47 (0.06)*
I connect with my own suffering without judging myself.	0.72 (0.02)*	0.79 (0.03)*
When I'm upset, I can let the emotions be there without feeling overwhelmed.	0.57 (0.03)*	0.63 (0.05)*
I connect with my own distress without letting it overwhelm me.	0.61 (0.03)*	0.65 (0.04)*
Acting or being motivated to act to alleviate suffering		
When I'm upset, I try to do what's best for myself.	0.86 (0.01)*	0.83 (0.02)*
When I'm going through a difficult time, I try to look after myself.	0.91 (0.01)*	0.87 (0.02)*
I try to make myself feel better when I'm distressed, even if I can't do anything about the cause.	0.76 (0.02)*	0.80 (0.02)*
When I'm upset, I do my best to take care of myself.	0.88 (0.01)*	0.85 (0.02)*

Standard errors are given in parentheses. * $p < .001$.

Table 7.
Correlation coefficients between total scale and subscale scores on the SOCS-O and SOCS-S using available data from 371 students (Stage 4)

	SOCS-O	1	2	3	4	5	SOCS-S	6	7	8	9	10
Sussex-Oxford Compassion for Others Scale (SOCS-O)	-											
1. Recognising suffering	.72***	-										
2. Understanding the universality of suffering	.62***	.25***	-									
3. Feeling for the person suffering	.79***	.47***	.31***	-								
4. Tolerating uncomfortable feelings	.77***	.46***	.37***	.49***	-							
5. Acting or being motivated to act to alleviate suffering	.83***	.49***	.32***	.69***	.61***	-						
Sussex-Oxford Compassion for the Self Scale (SOCS-S)	.34***	.14**	.44***	.18***	.29***	.19***	-					
6. Recognising suffering	.32***	.26***	.29***	.17**	.22***	.23***	.58***	-				
7. Understanding the universality of suffering	.43***	.13*	.74***	.20***	.29***	.22***	.52***	.32***	-			
8. Feeling for the person suffering	.21***	.08	.25***	.16**	.18**	.12*	.87***	.34***	.22***	-		
9. Tolerating uncomfortable feelings	.14**	.05	.19***	.05	.23***	.01	.79***	.25***	.20***	.72***	-	
10. Acting or being motivated to act to alleviate suffering	.17**	.03	.21***	.10	.18**	.13*	.86***	.29***	.24***	.82***	.67***	-

Note. *$p < .05$, **$p < .01$, ***$p < .001$.

Sussex-Oxford Compassion for Others Scale (SOCS-O)

Instructions

Below are statements describing how you might relate to **other people**. Please indicate how true the following statements are of you using the 5-point response scale (1 = Not at all true, 2 = Rarely true, 3 = Sometimes true, 4 = Often true, 5 = Always true). For example, if you think that a statement is often true of you, circle '4'.

Note: In the below items, generic terms (e.g., 'upset', 'distress', 'suffering', struggling') are used to cover a range of unpleasant emotions, such as sadness, fear, anger, frustration, guilt, shame, etc.

Please provide an answer for each statement.

	Not at all true	Rarely true	Sometimes true	Often true	Always true
1. I recognise when other people are feeling distressed without them having to tell me.	1	2	3	4	5
2. I understand that everyone experiences suffering at some point in their lives.	1	2	3	4	5
3. When someone is going through a difficult time, I feel kindly towards them.	1	2	3	4	5
4. When someone else is upset, I try to stay open to their feelings rather than avoid them.	1	2	3	4	5
5. When others are struggling, I try to do things that would be helpful.	1	2	3	4	5
6. I notice when others are feeling distressed.	1	2	3	4	5
7. I understand that feeling upset at times is part of human nature.	1	2	3	4	5
8. When I hear about bad things happening to other people, I feel concern for their wellbeing.	1	2	3	4	5
9. I stay with and listen to other people when they're upset even if it's hard to bear.	1	2	3	4	5
10. When someone is going through a difficult time, I try to look after them.	1	2	3	4	5
11. I'm quick to notice early signs of distress in others.	1	2	3	4	5
12. Like me, I know that other people also experience struggles in life.	1	2	3	4	5

	Not at all true	Rarely true	Sometimes true	Often true	Always true
13. When someone is upset, I try to tune in to how they're feeling.	1	2	3	4	5
14. I connect with the suffering of others without judging them.	1	2	3	4	5
15. When I see someone in need, I try to do what's best for them.	1	2	3	4	5
16. I recognise signs of suffering in others.	1	2	3	4	5
17. I know that we can all feel upset at times when we are wronged.	1	2	3	4	5
18. I'm sensitive to other people's distress.	1	2	3	4	5
19. When someone else is upset, I can be there for them without feeling overwhelmed by their distress.	1	2	3	4	5
20. When I see that someone is upset, I do my best to take care of them.	1	2	3	4	5

Scoring Information

Recognising suffering items: 1, 6, 11, 16.
Understanding the universality of suffering items: 2, 7, 12, 17.
Feeling for the person suffering items: 3, 8, 13, 18.
Tolerating uncomfortable feelings items: 4, 9, 14, 19.
Acting or being motivated to act to alleviate suffering items: 5, 10, 15, 20.

Sussex-Oxford Compassion for the Self Scale (SOCS-S)

Instructions

Below are statements describing how you might relate to **yourself**. Please indicate how true the following statements are of you using the 5-point response scale (1 = Not at all true, 2 = Rarely true, 3 = Sometimes true, 4 = Often true, 5 = Always true). For example, if you think that a statement is often true of you, circle '4'.

Note: In the below items, generic terms (e.g., 'upset', 'distress', 'suffering', struggling') are used to cover a range of unpleasant emotions, such as sadness, fear, anger, frustration, guilt, shame, etc.

Please provide an answer for each statement.

	Not at all true	Rarely true	Sometimes true	Often true	Always true
1. I'm good at recognising when I'm feeling distressed.	1	2	3	4	5
2. I understand that everyone experiences suffering at some point in their lives.	1	2	3	4	5
3. When I'm going through a difficult time, I feel kindly towards myself.	1	2	3	4	5
4. When I'm upset, I try to stay open to my feelings rather than avoid them.	1	2	3	4	5
5. I try to make myself feel better when I'm distressed, even if I can't do anything about the cause.	1	2	3	4	5
6. I notice when I'm feeling distressed.	1	2	3	4	5
7. I understand that feeling upset at times is part of human nature.	1	2	3	4	5
8. When bad things happen to me, I feel caring towards myself.	1	2	3	4	5
9. I connect with my own distress without letting it overwhelm me.	1	2	3	4	5
10. When I'm going through a difficult time, I try to look after myself.	1	2	3	4	5
11. I'm quick to notice early signs of distress in myself.	1	2	3	4	5
12. Like me, I know that other people also experience struggles in life.	1	2	3	4	5

	Not at all true	Rarely true	Sometimes true	Often true	Always true
13. When I'm upset, I try to tune in to how I'm feeling.	1	2	3	4	5
14. I connect with my own suffering without judging myself.	1	2	3	4	5
15. When I'm upset, I try to do what's best for myself.	1	2	3	4	5
16. I recognise signs of suffering in myself.	1	2	3	4	5
17. I know that we can all feel distressed when things don't go well in our lives.	1	2	3	4	5
18. Even when I'm disappointed with myself, I can feel warmly towards myself when I'm in distress.	1	2	3	4	5
19. When I'm upset, I can let the emotions be there without feeling overwhelmed.	1	2	3	4	5
20. When I'm upset, I do my best to take care of myself.	1	2	3	4	5

Scoring Information

Recognising suffering items: 1, 6, 11, 16.
Understanding the universality of suffering items: 2, 7, 12, 17.
Feeling for the person suffering items: 3, 8, 13, 18.
Tolerating uncomfortable feelings items: 4, 9, 14, 19.
Acting or being motivated to act to alleviate suffering items: 5, 10, 15, 20.

Appendix F: Certificate of Ethical Approval for Chapter 7 (Stages 1 and 2)

University of Sussex

Certificate of Approval	
Reference Number	ER/JG252/10
Title Of Project	Developing a new self-report measure of compassion: Stages 1 and 2
Principal Investigator (PI):	Clara Strauss
Student	Jenny Gu
Collaborators	
Duration Of Approval	6 months
Expected Start Date	31-Mar-2016
Date Of Approval	13-Apr-2016
Approval Expiry Date	30-Sep-2016
Approved By	David Reby
Name of Authorised Signatory	
Date	13-Apr-2016

*NB. If the actual project start date is delayed beyond 12 months of the expected start date, this Certificate of Approval will lapse and the project will need to be reviewed again to take account of changed circumstances such as legislation, sponsor requirements and University procedures.

Please note and follow the requirements for approved submissions:

Amendments to protocol

* Any changes or amendments to approved protocols must be submitted to the C-REC for authorisation prior to implementation.

Feedback regarding the status and conduct of approved projects

* Any incidents with ethical implications that occur during the implementation of the project must be reported immediately to the Chair of the C-REC.

Feedback regarding any adverse and unexpected events

* Any adverse (undesirable and unintended) and unexpected events that occur during the implementation of the project must be reported to the Chair of the Social Sciences C-REC. In the event of a serious adverse event, research must be stopped immediately and the Chair alerted within 24 hours of the occurrence.

For Life Sciences and Psychology projects

* The principal investigator is required to provide a brief annual written statement to the committee, indicating the status and conduct of the approved project. These reports will be reviewed at the annual meeting of the committee. A statement by the PI to the C-REC indicating the status and conduct of the approved project will be required on the Approval Expiration Date as stated above.

Appendix F: Certificates of Ethical Approval for Chapter 7 (Stages 3 and 4)

University of Sussex

Certificate of Approval	
Reference Number	ER/JG252/11
Title Of Project	Developing a new self-report measure of compassion: Stages 3 and 4
Principal Investigator (PI):	Clara Strauss
Student	Jenny Gu
Collaborators	
Duration Of Approval	9 months
Expected Start Date	04-Jan-2016
Date Of Approval	29-Nov-2016
Approval Expiry Date	30-Sep-2017
Approved By	David Reby
Name of Authorised Signatory	
Date	29-Nov-2016

*NB. If the actual project start date is delayed beyond 12 months of the expected start date, this Certificate of Approval will lapse and the project will need to be reviewed again to take account of changed circumstances such as legislation, sponsor requirements and University procedures.

Please note and follow the requirements for approved submissions:

Amendments to protocol

* Any changes or amendments to approved protocols must be submitted to the C-REC for authorisation prior to implementation.

Feedback regarding the status and conduct of approved projects

* Any incidents with ethical implications that occur during the implementation of the project must be reported immediately to the Chair of the C-REC.

Feedback regarding any adverse and unexpected events

* Any adverse (undesirable and unintended) and unexpected events that occur during the implementation of the project must be reported to the Chair of the Social Sciences C-REC. In the event of a serious adverse event, research must be stopped immediately and the Chair alerted within 24 hours of the occurrence.

For Life Sciences and Psychology projects

* The principal investigator is required to provide a brief annual written statement to the committee, indicating the status and conduct of the approved project. These reports will be reviewed at the annual meeting of the committee. A statement by the PI to the C-REC indicating the status and conduct of the approved project will be required on the Approval Expiration Date as stated above.

Certificate of Approval	
Reference Number	ER/JG252/12
Title Of Project	Developing a new self-report measure of compassion: Validation in a student sample
Principal Investigator (PI):	Clara Strauss
Student	Jenny Gu
Collaborators	
Duration Of Approval	4 months
Expected Start Date	25-Sep-2017
Date Of Approval	21-Sep-2017
Approval Expiry Date	29-Dec-2017
Approved By	David Reby
Name of Authorised Signatory	Anna Hobbs
Date	21-Sep-2017

*NB. If the actual project start date is delayed beyond 12 months of the expected start date, this Certificate of Approval will lapse and the project will need to be reviewed again to take account of changed circumstances such as legislation, sponsor requirements and University procedures.

Please note and follow the requirements for approved submissions:

Amendments to protocol

* Any changes or amendments to approved protocols must be submitted to the C-REC for authorisation prior to implementation.

Feedback regarding the status and conduct of approved projects

* Any incidents with ethical implications that occur during the implementation of the project must be reported immediately to the Chair of the C-REC.

Feedback regarding any adverse(1) and unexpected events(2)

* Any adverse (undesirable and unintended) and unexpected events that occur during the implementation of the project must be reported to the Chair of the Science and Technology C-REC. In the event of a serious adverse event, research must be stopped immediately and the Chair alerted within 24 hours of the occurrence.

Monitoring of Approved studies

The University may undertake periodic monitoring of approved studies. Researchers will be requested to report on the outcomes of research activity in relation to approvals that were granted (full applications and amendments).

Research Standards

Failure to conduct University research in alignment with the Code of Practice for Research may be investigated under the Procedure for the Investigation of Allegations of Misconduct in Research or other appropriate internal mechanisms (3). Any queries can be addressed to the Research Governance Office: rgoffice@sussex.ac.uk

(1) An "adverse event" is one that occurs during the course of a research protocol that either causes physical or psychological harm, or increases the risk of physical or psychological harm, or results in a loss of privacy and/or confidentiality to research participant or others.

(2) An "unexpected event" is an occurrence or situation during the course of a research project that was a) harmful to a participant taking part in the research, or b) increased the probability of harm to participants taking part in the research.

(3) http://www.sussex.ac.uk/staff/research/rqi/policy/research-policy

NHS
Health Research Authority

Miss Jenny Gu
PhD student
University of Sussex
Pevensey 1 2D5
Falmer
BN1 9RH

Email: hra.approval@nhs.net

05 January 2017

Dear Miss Gu,

Letter of HRA Approval

Study title:	**Developing and validating a new self-report measure of compassion**
IRAS project ID:	**216006**
REC reference:	**17/HRA/0023**
Sponsor	**University of Sussex**

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

Participation of NHS Organisations in England

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

Appendix B provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

Further information on funding, HR processes, and compliance with HRA criteria and standards is also provided.